

## TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

## NOTIFICATION D'ELECTION

(règle 61.2 du PCT)

Expéditeur: le BUREAU INTERNATIONAL

Destinataire:

Commissioner  
 US Department of Commerce  
 United States Patent and Trademark  
 Office, PCT  
 2011 South Clark Place Room  
 CP2/5C24  
 Arlington, VA 22202  
 ETATS-UNIS D'AMERIQUE  
 en sa qualité d'office élu

Date d'expédition (jour/mois/année) 26 février 2001 (26.02.01)	
Demande internationale no PCT/FR00/01904	Référence du dossier du déposant ou du mandataire BCT000059/BLO
Date du dépôt international (jour/mois/année) 04 juillet 2000 (04.07.00)	Date de priorité (jour/mois/année) 05 juillet 1999 (05.07.99)
Déposant CAPMAN, François etc	

1. L'office désigné est avisé de son élection qui a été faite:



dans la demande d'examen préliminaire international présentée à l'administration chargée de l'examen préliminaire international le:

10 janvier 2001 (10.01.01)



dans une déclaration visant une élection ultérieure déposée auprès du Bureau international le:

2. L'élection



a été faite



n'a pas été faite

avant l'expiration d'un délai de 19 mois à compter de la date de priorité ou, lorsque la règle 32 s'applique, dans le délai visé à la règle 32.2b).

Bureau international de l'OMPI 34, chemin des Colombettes 1211 Genève 20, Suisse no de télécopieur: (41-22) 740.14.35	Fonctionnaire autorisé Kiwa Mpay no de téléphone: (41-22) 338.83.38
--	---

THIS PAGE BLANK (USPTO)

## RAPPORT DE RECHERCHE INTERNATIONALE

(article 18 et règles 43 et 44 du PCT)

Référence du dossier du déposant ou du mandataire <b>BCT000059/BLO</b>	<b>POUR SUITE</b> voir la notification de transmission du rapport de recherche internationale (formulaire PCT/ISA/220) et, le cas échéant, le point 5 ci-après <b>A DONNER</b>	
Demande internationale n° <b>PCT/FR 00/ 01904</b>	Date du dépôt international(jour/mois/année) <b>04/07/2000</b>	(Date de priorité (la plus ancienne) (jour/mois/année) <b>05/07/1999</b>
Déposant  <b>MATRA NORTEL COMMUNICATIONS</b>		

Le présent rapport de recherche internationale, établi par l'administration chargée de la recherche internationale, est transmis au déposant conformément à l'article 18. Une copie en est transmise au Bureau international.

Ce rapport de recherche internationale comprend 3 feuilles.



Il est aussi accompagné d'une copie de chaque document relatif à l'état de la technique qui y est cité.

## 1. Base du rapport

- a. En ce qui concerne la langue, la recherche internationale a été effectuée sur la base de la demande internationale dans la langue dans laquelle elle a été déposée, sauf indication contraire donnée sous le même point.



la recherche internationale a été effectuée sur la base d'une traduction de la demande internationale remise à l'administration.

- b. En ce qui concerne les séquences de nucléotides ou d'acides aminés divulguées dans la demande internationale (le cas échéant), la recherche internationale a été effectuée sur la base du listage des séquences :



contenu dans la demande internationale, sous forme écrite.



déposée avec la demande internationale, sous forme déchiffrable par ordinateur.



remis ultérieurement à l'administration, sous forme écrite.



remis ultérieurement à l'administration, sous forme déchiffrable par ordinateur.



La déclaration, selon laquelle le listage des séquences présenté par écrit et fourni ultérieurement ne vas pas au-delà de la divulgation faite dans la demande telle que déposée, a été fournie.



La déclaration, selon laquelle les informations enregistrées sous forme déchiffrable par ordinateur sont identiques à celles du listage des séquences présenté par écrit, a été fournie.

2. ☐ Il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (voir le cadre I).

3. ☐ Il y a absence d'unité de l'invention (voir le cadre II).

## 4. En ce qui concerne le titre,



le texte est approuvé tel qu'il a été remis par le déposant.



Le texte a été établi par l'administration et a la teneur suivante:

## 5. En ce qui concerne l'abrégé,



le texte est approuvé tel qu'il a été remis par le déposant.



le texte (reproduit dans le cadre III) a été établi par l'administration conformément à la règle 38.2b). Le déposant peut présenter des observations à l'administration dans un délai d'un mois à compter de la date d'expédition du présent rapport de recherche internationale.

## 6. La figure des dessins à publier avec l'abrégé est la Figure n°



suggérée par le déposant.



parce que le déposant n'a pas suggéré de figure.



parce que cette figure caractérise mieux l'invention.

15



Aucune des figures n'est à publier.

**THIS PAGE BLANK (USPTO)**

# RAPPORT DE RECHERCHE INTERNATIONALE

Dei. le internationale No

PCT/FR 00/01904

### A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 7 G10L19/02

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 G10L

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, WPI Data, INSPEC

### C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	US 5 878 388 A (NISHIGUCHI MASAYUKI ET AL) 2 mars 1999 (1999-03-02)	1,4,6,7, 16
Y	figures 2A,2B colonne 5, ligne 63 -colonne 6, ligne 65 colonne 9, ligne 27-39 colonne 10, ligne 24-44 colonne 11, ligne 1-5 colonne 11, ligne 36-52 colonne 12, ligne 42-52 colonne 33, ligne 35 -colonne 34, ligne 8	5,8,9
Y	US 5 911 130 A (SHIMIZU ET AL) 8 juin 1999 (1999-06-08)	8,9
A	colonne 4, ligne 48-65; figure 7 colonne 7, ligne 49 -colonne 8, ligne 37	11,12
	-/-	

**X**

Voir la suite du cadre C pour la fin de la liste des documents

☒

Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

\*A\* document définissant l'état général de la technique, non considéré comme particulièrement pertinent

**"E"** document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

\*O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens.

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

<sup>T</sup> document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

**"g" document qui fait partie de la même famille de brevets**

Date à laquelle la recherche internationale a été effectivement achevée

6 septembre 2000

Date d'expédition du présent rapport de recherche internationale

14/09/2000

\_\_\_\_\_  
Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Quélavoine, R

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	<p>DATABASE INSPEC 'en ligne!  INSTITUTE OF ELECTRICAL ENGINEERS,  STEVENAGE, GB;  LAGLER A ET AL: "Real-time fixed-point  DSP-implementation of spectral  subtraction algorithm for speech  enhancement in noisy environment"  Database accession no. 6604117  XP002139930  abrégé  &amp; 9TH INTERNATIONAL CZECH - SLOVAK  SCIENTIFIC CONFERENCE. RADIOELEKTRONIKA  99. CONFERENCE PROCEEDINGS, PROCEEDINGS OF  9TH INTERNATIONAL CZECH - SLOVAK  SCIENTIFIC CONFERENCE. RADIOELEKTRONIKA  99, BRNO, CZECH REPUBLIC, 27-28 APRIL  1999,  pages 186-189,  1999, Brno, Czech Republic, Brno Univ.  Technol, Czech Republic  ISBN: 80-214-1327-1</p>	5
A	<p>AHMADI S ET AL: "New techniques for  sinusoidal coding of speech at 2400 bps"  CONFERENCE RECORD OF THIRTIETH ASILOMAR  CONFERENCE ON SIGNALS, SYSTEMS AND  COMPUTERS (CAT. NO.96CB36004), CONFERENCE  RECORD OF THE THIRTIETH ASILOMAR  CONFERENCE ON SIGNALS, SYSTEMS AND  COMPUTERS, PACIFIC GROVE, CA, USA, 3-6  NOV. 1996,  pages 770-774 vol.1, XP002138769  1997, Los Alamitos, CA, USA, IEEE Comput.  Soc. Press, USA  ISBN: 0-8186-7646-9  page 773, colonne de droite, ligne 10-30</p>	10
A	<p>EP 0 893 791 A (DIGITAL VOICE SYSTEMS INC)  27 janvier 1999 (1999-01-27)</p> <p>abrégé  page 2, colonne 48 -page 3, colonne 1  page 6, ligne 21-29  page 7, ligne 36-43</p>	1,2,4,6, 7,10,13, 16

# INTERNATIONAL SEARCH REPORT

Int'l Application No

PCT/FR 00/01904

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 878 388 A (NISHIGUCHI MASAYUKI ET AL) 2 March 1999 (1999-03-02)	1,4,6,7,16
Y	figures 2A,2B column 5, line 63 -column 6, line 65 column 9, line 27-39 column 10, line 24-44 column 11, line 1-5 column 11, line 36-52 column 12, line 42-52 column 33, line 35 -column 34, line 8	5,8,9
Y	US 5 911 130 A (SHIMIZU ET AL) 8 June 1999 (1999-06-08)	8,9
A	column 4, line 48-65; figure 7 column 7, line 49 -column 8, line 37	11,12
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*S\* document member of the same patent family

Date of the actual completion of the international search

6 September 2000

Date of mailing of the international search report

14/09/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

# INTERNATIONAL SEARCH REPORT

Int'l Application No

PCT/FR 00/01904

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>DATABASE INSPEC 'Online!  INSTITUTE OF ELECTRICAL ENGINEERS,  STEVENAGE, GB;  LAGLER A ET AL: "Real-time fixed-point  DSP-implementation of spectral  subtraction algorithm for speech  enhancement in noisy environment"  Database accession no. 6604117  XP002139930  abstract  &amp; 9TH INTERNATIONAL CZECH - SLOVAK  SCIENTIFIC CONFERENCE. RADIOELEKTRONIKA  99. CONFERENCE PROCEEDINGS, PROCEEDINGS OF  9TH INTERNATIONAL CZECH - SLOVAK  SCIENTIFIC CONFERENCE. RADIOELEKTRONIKA  99, BRNO, CZECH REPUBLIC, 27-28 APRIL  1999,  pages 186-189,  1999, Brno, Czech Republic, Brno Univ.  Technol, Czech Republic  ISBN: 80-214-1327-1</p>	5
A	<p>AHMADI S ET AL: "New techniques for  sinusoidal coding of speech at 2400 bps"  CONFERENCE RECORD OF THIRTIETH ASILOMAR  CONFERENCE ON SIGNALS, SYSTEMS AND  COMPUTERS (CAT. NO.96CB36004), CONFERENCE  RECORD OF THE THIRTIETH ASILOMAR  CONFERENCE ON SIGNALS, SYSTEMS AND  COMPUTERS, PACIFIC GROVE, CA, USA, 3-6  NOV. 1996,  pages 770-774 vol.1, XP002138769  1997, Los Alamitos, CA, USA, IEEE Comput.  Soc. Press, USA  ISBN: 0-8186-7646-9  page 773, right-hand column, line 10-30</p>	10
A	<p>EP 0 893 791 A (DIGITAL VOICE SYSTEMS INC)  27 January 1999 (1999-01-27)   abstract  page 2, column 48 -page 3, column 1  page 6, line 21-29  page 7, line 36-43</p>	1,2,4,6, 7,10,13, 16



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PC 00/01904

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
US 5878388	A	02-03-1999	JP	5265487 A		15-10-1993
			JP	5297894 A		12-11-1993
			EP	0590155 A		06-04-1994
			WO	9319459 A		30-09-1993
			US	5765127 A		09-06-1998
			US	5960388 A		28-09-1999
<hr/>						
US 5911130	A	08-06-1999	JP	8330971 A		13-12-1996
<hr/>						
EP 0893791	A	27-01-1999	US	5226084 A		06-07-1993
			AU	657508 B		16-03-1995
			AU	9147091 A		08-07-1992
			CA	2096425 A		06-06-1992
			DE	69132013 D		06-04-2000
			EP	0560931 A		22-09-1993
			JP	6507252 T		11-08-1994
			WO	9210830 A		25-06-1992
			US	5630011 A		13-05-1997
			US	5491772 A		13-02-1996
			US	5247579 A		21-09-1993
<hr/>						

THIS PAGE BLANK (USPTO)

# TRAITE DE COOPERATION EN MATIERE DE BREVETS

Expéditeur: L'ADMINISTRATION CHARGÉE DE  
L'EXAMEN PRELIMINAIRE INTERNATIONAL

Destinataire:

LOISEL, Bertrand  
CABINET PLASSERAUD  
84, rue d'Amsterdam  
75440 Paris Cedex 09  
FRANCE

REÇU LE

28 SEP. 2001

Cbt Plasseraud

PCT

NOTIFICATION DE TRANSMISSION DU  
RAPPORT D'EXAMEN PRELIMINAIRE  
INTERNATIONAL  
(règle 71.1 du PCT)

Date d'expédition  
(jour/mois/année) 26.09.2001

Référence du dossier du déposant ou du mandataire  
BCT000059/BLO

**NOTIFICATION IMPORTANTE**

Demande internationale No.  
PCT/FR00/01904

Date du dépôt international (jour/mois/année)  
04/07/2000

Date de priorité (jour/mois/année)  
05/07/1999

Déposant  
MATRA NORTEL COMMUNICATIONS et al.

1. Il est notifié au déposant que l'administration chargée de l'examen préliminaire international a établi le rapport d'examen préliminaire international pour la demande internationale et le lui transmet ci-joint, accompagné, le cas échéant, de ces annexes.

2. Une copie du présent rapport et, le cas échéant, de ses annexes est transmise au Bureau international pour communication à tous les offices élus.

3. Si tel ou tel office élu l'exige, le Bureau international établira une traduction en langue anglaise du rapport (à l'exclusion des annexes de celui-ci) et la transmettra aux offices intéressés.

#### 4. RAPPEL

Pour aborder la phase nationale auprès de chaque office élu, le déposant doit accomplir certains actes (dépôt de traduction et paiement des taxes nationales) dans le délai de 30 mois à compter de la date de priorité (ou plus tard pour ce qui concerne certains offices) (article 39.1) (voir aussi le rappel envoyé par le Bureau international dans le formulaire PCT/IB/301).

Lorsqu'une traduction de la demande internationale doit être remise à un office élu, elle doit comporter la traduction de toute annexe du rapport d'examen préliminaire international. Il appartient au déposant d'établir la traduction en question et de la remettre directement à chaque office élu intéressé.

Pour plus de précisions en ce qui concerne les délais applicables et les exigences des offices élus, voir le Volume II du Guide du déposant du PCT.

Nom et adresse postale de l'administration chargée de l'examen  
préliminaire international



Office européen des brevets  
D-80298 Munich  
Tél. +49 89 2399 - 0 Tx: 523656 epmu d  
Fax: +49 89 2399 - 4465

Fonctionnaire autorisé

Benigar, M

Tél. +49 89 2399-2996



**THIS PAGE BLANK (USPTO)**

PCT

RAPPORT D'EXAMEN PRELIMINAIRE INTERNATIONAL



(article 36 et règle 70 du PCT)

Référence du dossier du déposant ou du mandataire BCT000059/BLO	POUR SUITE A DONNER voir la notification de transmission du rapport d'examen préliminaire international (formulaire PCT/IPEA/416)	
Demande internationale n° PCT/FR00/01904	Date du dépôt international (jour/mois/année) 04/07/2000	Date de priorité (jour/mois/année) 05/07/1999
Classification internationale des brevets (CIB) ou à la fois classification nationale et CIB G10L19/02		
Déposant MATRA NORTEL COMMUNICATIONS et al.		

1. Le présent rapport d'examen préliminaire international, établi par l'administration chargée de l'examen préliminaire international, est transmis au déposant conformément à l'article 36.
2. Ce RAPPORT comprend 9 feuilles, y compris la présente feuille de couverture.
  - ☐ Il est accompagné d'ANNEXES, c'est-à-dire de feuilles de la description, des revendications ou des dessins qui ont été modifiées et qui servent de base au présent rapport ou de feuilles contenant des rectifications faites auprès de l'administration chargée de l'examen préliminaire international (voir la règle 70.16 et l'instruction 607 des Instructions administratives du PCT).

Ces annexes comprennent feuilles.

3. Le présent rapport contient des indications relatives aux points suivants:
  - I ☒ Base du rapport
  - II ☐ Priorité
  - III ☐ Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle
  - IV ☐ Absence d'unité de l'invention
  - V ☒ Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration
  - VI ☐ Certains documents cités
  - VII ☒ Irrégularités dans la demande internationale
  - VIII ☒ Observations relatives à la demande internationale

Date de présentation de la demande d'examen préliminaire internationale 10/01/2001	Date d'achèvement du présent rapport 26.09.2001
Nom et adresse postale de l'administration chargée de l'examen préliminaire international:  Office européen des brevets D-80298 Munich Tél. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Fonctionnaire autorisé Bourdier, R N° de téléphone +49 89 2399 2130 

**THIS PAGE BLANK (USPTO)**

# RAPPORT D'EXAMEN PRÉLIMINAIRE INTERNATIONAL

Demande internationale n° PCT/FR00/01904

## I. Base du rapport

1. En ce qui concerne les **éléments** de la demande internationale (*les feuilles de remplacement qui ont été remises à l'office récepteur en réponse à une invitation faite conformément à l'article 14 sont considérées dans le présent rapport comme "initialement déposées" et ne sont pas jointes en annexe au rapport puisqu'elles ne contiennent pas de modifications (règles 70.16 et 70.17)*):

### Description, pages:

1-30                      version initiale

### Revendications, N°:

1-16                      version initiale

### Dessins, feuilles:

1/11-11/11              version initiale

2. En ce qui concerne la **langue**, tous les éléments indiqués ci-dessus étaient à la disposition de l'administration ou lui ont été remis dans la langue dans laquelle la demande internationale a été déposée, sauf indication contraire donnée sous ce point.

Ces éléments étaient à la disposition de l'administration ou lui ont été remis dans la langue suivante: , qui est :

- ☐ la langue d'une traduction remise aux fins de la recherche internationale (selon la règle 23.1(b)).
- ☐ la langue de publication de la demande internationale (selon la règle 48.3(b)).
- ☐ la langue de la traduction remise aux fins de l'examen préliminaire internationale (selon la règle 55.2 ou 55.3).

3. En ce qui concerne les **séquences de nucléotides ou d'acide aminés** divulguées dans la demande internationale (le cas échéant), l'examen préliminaire internationale a été effectué sur la base du listage des séquences :

- ☐ contenu dans la demande internationale, sous forme écrite.
- ☐ déposé avec la demande internationale, sous forme déchiffrable par ordinateur.
- ☐ remis ultérieurement à l'administration, sous forme écrite.
- ☐ remis ultérieurement à l'administration, sous forme déchiffrable par ordinateur.
- ☐ La déclaration, selon laquelle le listage des séquences par écrit et fourni ultérieurement ne va pas au-delà de la divulgation faite dans la demande telle que déposée, a été fournie.
- ☐ La déclaration, selon laquelle les informations enregistrées sous déchiffrable par ordinateur sont identiques à celles du listage des séquences Présenté par écrit, a été fournie.

4. Les modifications ont entraîné l'annulation :

**THIS PAGE BLANK (USPTO)**



**RAPPORT D'EXAMEN  
PRÉLIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR00/01904

- ☐ de la description, pages :  
☐ des revendications, n°s :  
☐ des dessins, feuilles :

5. ☐ Le présent rapport a été formulé abstraction faite (de certaines) des modifications, qui ont été considérées comme allant au-delà de l'exposé de l'invention tel qu'il a été déposé, comme il est indiqué ci-après (règle 70.2(c)) :

*(Toute feuille de remplacement comportant des modifications de cette nature doit être indiquée au point 1 et annexée au présent rapport)*

6. Observations complémentaires, le cas échéant :

**V. Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration**

**1. Déclaration**

Nouveauté	Oui : Revendications 1-5,7-15 Non : Revendications 6,16
Activité inventive	Oui : Revendications 2,3,13,14 Non : Revendications 1,4,5,7-12,15
Possibilité d'application industrielle	Oui : Revendications 1-16 Non : Revendications

**2. Citations et explications  
voir feuille séparée**

**VII. Irrégularités dans la demande internationale**

Les irrégularités suivantes, concernant la forme ou le contenu de la demande internationale, ont été constatées :  
**voir feuille séparée**

**VIII. Observations relatives à la demande internationale**

Les observations suivantes sont faites au sujet de la clarté des revendications, de la description et des dessins et de la question de savoir si les revendications se fondent entièrement sur la description :  
**voir feuille séparée**

**THIS PAGE BLANK (USPTO)**

**CONCERNANT LA SECTION V**

1. L'état de la technique le plus proche pour la revendication 1 est considéré comme étant US-A-5 878 388, que l'on désignera par D1.

D1 divulgue une procédé d'analyse d'un signal audio (signal de parole, colonne 5, ligne 63) traité par trames successives de N échantillons, pondérées par une fenêtre de Hamming (colonne 6, lignes 12-14), dans lequel on calcule un spectre du signal audio pondéré par transformation fréquentielle (colonne 6, lignes 59-65), et on traite le spectre pour délivrer des paramètres de synthèse (colonne 9, lignes 40-47 et colonne 7, lignes 1-3), les trames successives pour lesquels on délivre des jeux de paramètres présentant des recouvrements mutuels de moins de  $N/2$  échantillons (colonne 6, lignes 8-11: longueur de la fenêtre=256, recouvrement=96).

La caractéristique selon laquelle on délivre en alternance des jeux complets de paramètres et des jeux incomplets est nouvelle par rapport à D1.

2. Le problème technique qu'on résout en introduisant la caractéristique nouvelle ci-dessus est bien entendu celui de limiter le débit de codage.

Le document D2 = AHMADI S ET AL: 'New techniques for sinusoidal coding of speech at 2400 bps' CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), CONFERENCE RECORD OF THE THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA, divulgue une technique de codage de parole délivrant des paramètres spectraux sur des trames de longueur fixe (page 770, "Introduction", colonne de gauche). Afin de limiter la bande passante du codeur, ce document propose de limiter l'envoi de paramètres de codage à une trame sur 2, les paramètres manquants étant ré-estimés par interpolation (page 773, chapitre 5, lignes 1-16). Toutes les 2 trames dans D2, le jeu de paramètres est donc incomplet puisqu'il est vide.

**THIS PAGE BLANK (USPTO)**

L'objet de la revendication 1 résulte donc de l'introduction de l'enseignement de D2 concernant la réduction de débit dans le procédé connu de D1. Une telle juxtaposition de caractéristiques est considérée comme évidente et donc manquant de satisfaire au critère d'activité inventive.

L'argument consistant à considérer qu'un recouvrement des fenêtres inférieur à  $N/2$  est préjudiciable à l'envoi d'une trame sur deux ne peut être retenu pour justifier l'activité inventive. En effet, il est parfaitement évident que plus l'espacement des fenêtres est important, plus les paramètres spectraux changent à chaque fenêtre, et donc plus le codage perd en qualité, surtout si l'on cherche en plus, comme dans D2, à interpoler une trame sur deux. Ainsi l'homme du métier choisira l'espacement des fenêtres en connection avec l'interpolation en connaissance de cause, selon des critères résultant d'un compromis entre baisse de débit et qualité de codage. En particulier rien ne l'empêche de choisir un type de codage comme revendiqué à la revendication 1, ce choix étant en outre, au vu de l'enseignement combiné de D1 et D2, le plus évident.

3. Les caractéristiques additionnelles de la revendication 4 sont également connues de D1 (voir passages cités plus haut). Donc la revendication 4 manque de satisfaire le critère d'activité inventive, tout comme la revendication 1.
4. Les caractéristiques additionnelles des revendications 2 et 3 sont nouvelles et considérées comme non évidente au vu de l'art antérieur, et donc impliquant une activité inventive.
5. La caractéristique additionnelle de la revendication 5 est nouvelle par rapport à D1 mais est commune dans la technique et n'a aucune incidence sur les caractéristiques de codage proprement dites liées aux trames et aux fenêtres. La juxtaposition est donc considérée comme non inventive.
6. N'importe quel processeur de signal serait capable d'exécuter le procédé des revendications 1 à 5. Vu le manque de clarté résultant de l'énoncé de la revendication 6 (voir section VIII), cette revendication n'est donc pas nouvelle.
7. L'état de la technique le plus proche pour la revendication 7 est considéré comme

**THIS PAGE BLANK (USPTO)**

étant le document D3 = "A weighted overlap-add method for short-time Fourier analysis synthesis", R.E. Crochiere, IEEE TRans. on ASSP, Vol ASSP 28, février 1980, pages 99-102, nouvellement cité.

D3 divulgue un procédé d'analyse synthèse par transformation de Fourier à court-terme (TFCT) et contient à ce titre toutes les caractéristiques du préambule de la revendication 7.

Selon le schéma de la figure 2, D3 applique aux séquences à court-terme resynthétisées une fenêtre de synthèse  $f(n)$  et additionne les portions de signal obtenues par addition et recouvrement. D3, en ce sens, est le plus général possible, puisqu'il ne donne aucune condition sur les fenêtres d'analyse et de synthèse. Il est bien connu cependant qu'un choix approprié de fenêtres d'analyse et de synthèse permet d'obtenir, par exemple, un signal synthétisé identique à l'original (voir D3, page 100, colonne de droite, paragraphe entre les équations (8) et (9)).

En supposant la partie caractérisante de la revendication 7 clarifiée (voir section VIII), cette partie revient à s'assurer que les portions de signaux pondérées par application d'une fenêtre d'analyse multiplicative, recouvrent à la synthèse leur amplitude de départ par division de la fenêtre d'analyse  $f_A$ , et que les signaux à court-termes ainsi obtenus soient pondérés sur l'espace de leur recouvrement par la fenêtre de synthèse  $f_S$  de manière à ce que la somme des facteurs de pondération lors de l'addition des signaux par recouvrement soit égale à une constante, identique aux portions ne représentant pas de recouvrement. De telles mesures découlent de manière évidente d'un document comme D3 et ne sont que la formalisation des conditions à réaliser sur la forme des fenêtres, voir le dessin parfaitement connu de l'homme du métier présent par exemple dans US-A-5 911 130, figure 7 et colonne 8, lignes 24-38.

On remarquera que dans la technique, le rôle de la fenêtre de synthèse est souvent occulté. Ainsi, dans le schéma classique d'analyse/synthèse par TFCT ou les fenêtres d'analyses triangulaires ou de Hanning se recouvrent de moitié de manière à avoir une somme constante égale à l'unité, les signaux resynthétisés sont additionnés par recouvrement identique, sans application de fenêtre de

**THIS PAGE BLANK (USPTO)**



synthèse. Il est cependant évident que cette dernière opération conduit au même résultat qu'une division par la fenêtre d'analyse (pour retrouver l'amplitude 1 sur la trame), suivie de l'application de fenêtre de synthèse égale à la fenêtre d'analyse de départ (pour pondérer de manière adéquate sur les recouvrements), comme il est mentionné à la revendication 7. Il n'y a donc rien de nouveau là-dedans, si ce n'est un formalisme rigoureux pour appuyer une pratique parfaitement commune.

8. Les caractéristiques additionnelles des revendications 8 et 9 ne sont que des exemples de fenêtres connues dans la technique, qui en soi n'ont rien d'inventif.

9. L'état de la technique le plus pertinent pour la revendication 10 est considéré comme étant D1.

Conformément à l'analyse du point 2 ci-dessus, les caractéristiques de la revendication 10 jusqu'à page 33, ligne 1, sont rendues évidentes par une combinaison de D1 et D2.

Les caractéristiques additionnelles de la page 33, lignes 2-8 ne font que reprendre sous une forme différente (décalage M/p plutôt que L) les conditions sur les fenêtres d'analyse et de synthèse de la revendication 7. Pour les raisons invoquées au point 7 ci-dessus, l'examineur considère ces caractéristiques additionnelles comme évidentes pour un homme du métier, de sorte que, même combinées au reste de la revendication 10, aucune activité inventive n'est reconnue.

10. L'opinion négative pour les revendications 8 et 9 s'applique également aux revendications 11 et 12.

11. L'opinion positive pour les revendications 2 et 3 s'applique également aux revendications 13 et 14.

12. Même si D1 et D2 ne concernent que des coefficients spectraux, l'extension à des coefficients cepstraux comme revendiqué à la revendication 15, connus en soi, ne semble avoir aucune incidence sur les principes de codage. Donc l'objet de la revendication 15 manque de satisfaire au critère d'activité inventive pour la même raison que la revendication 10.

**THIS PAGE BLANK (USPTO)**

13. La revendication 16 manque de nouveauté pour la même raison que la revendication 6 (voir aussi section VIII, points 1 et 2).
14. Toutes les revendications sont susceptibles d'activité industrielle.

#### CONCERNANT LA SECTION VII

1. Contrairement à ce qu'exige la règle 5.1 a) ii) PCT, la description n'indique pas l'état de la technique antérieure pertinent exposé dans les documents D1, D2 et D3 et ne cite pas ces documents.

#### CONCERNANT LA SECTION VIII

1. La revendication 6 mentionne un dispositif comprenant des "moyens pour exécuter" un procédé. Une telle tournure manque de clarté car elle doit être interprétée comme incluant des moyens simplement appropriés pour exécuter un tel procédé. En d'autres termes, les moyens revendiqués ne sont pas limités à ceux qui, effectivement, ont été programmés pour exécuter le procédé.

Afin de lever l'objection précédente, il conviendrait par exemple de remplacer l'expression "pour exécuter" par "programmés pour exécuter".

2. Une objection équivalente est formée contre la revendication 16.
3. La clause de la revendication 7 selon laquelle "on modifie chaque trame évaluée en lui appliquant un traitement correspondant à une division..." manque de clarté puisqu'en effet on se sait pas si le traitement revendiqué consiste réellement à effectuer une division suivi d'une multiplication, comme mentionné à la revendication 7, ou bien si le traitement, autre, conduit à un résultat identique.

**THIS PAGE BLANK (USPTO)**

Un exemple illustre cette ambiguïté: le procédé de synthèse par TFCT tel que décrit dans D3 peut s'appliquer à des fenêtres d'analyses et de synthèse rectangulaires d'amplitude égale à 1 de longueur N, décalées de N. L'analyse-synthèse consiste à mettre bout à bout les portions de signaux transformés, puis retransformés à l'identique. Dans ce cas, on n'a besoin d'appliquer aucun traitement aux trames évaluées, mais cette absence de traitement correspond à une division par la fenêtre d'analyse et à la multiplication par la fenêtre de synthèse, puisque ces opérations sont égales à l'identité.

---

**THIS PAGE BLANK (USPTO)**

(19) World Intellectual Property Organization

International Bureau

WIPO

(43) International publication date

11 January 2001 (11.01.2001)

PCT

(10) International publication number

WO 01/03116 A1

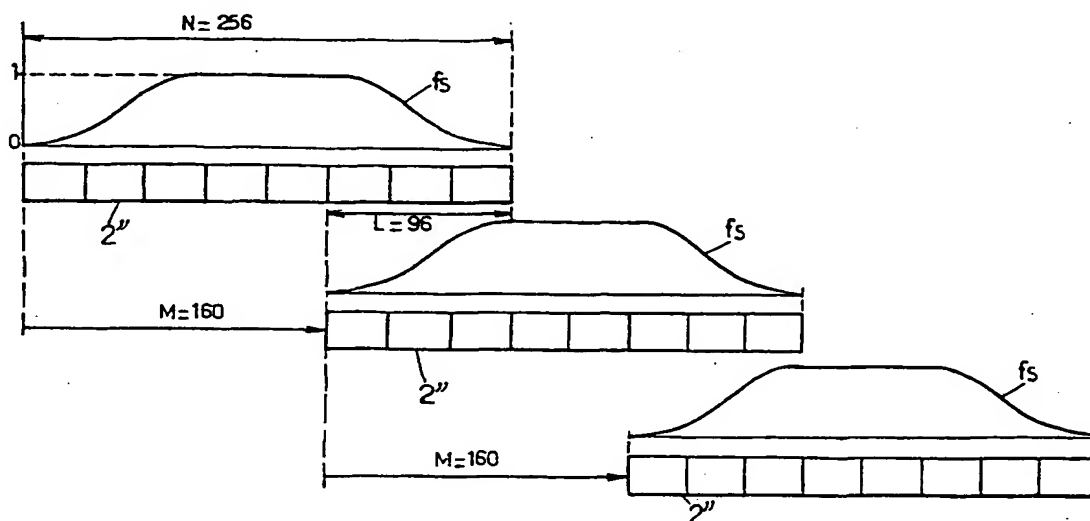
- (51) International patent classification<sup>7</sup>: G10L 19/02
- (21) International application number: PCT/FR00/01904
- (22) International filing date: 4 July 2000 (04.07.2000)
- (25) Language of filing: French
- (26) Language of publication: French
- (30) Data relating to the priority:  
99/08,638 5 July 1999 (05.07.1999) FR
- (71) Applicant (for all designated States except US):  
MATRA NORTEL COMMUNICATIONS [FR/FR]; 50, rue  
du Président Sadate, F-29100 Quimper (FR).
- (72) Inventors; and  
(75) Inventors/Applicants (US only): CAPMAN, François  
[FR/FR]; 47, rue des Etats Généraux, F-78000  
Versailles (FR). MURGIA, Carlo [IT/US]; 458 Saint  
Vincent, Irvine, CA 92618 (US).
- (74) Representatives: LOISEL, Bertrand etc.; Cabinet  
Plasseraud, 84, rue d'Amsterdam, F-75440 Paris Cedex  
09 (FR).
- (81) Designated states (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,  
CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD,  
GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,  
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,  
MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO,  
RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ,  
UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated states (regional): ARIPO Patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW),

[continued on next page]

As printed

(54) Title: METHODS AND DEVICE FOR AUDIO ANALYSIS AND SYNTHESIS

(54) Titre: PROCÉDES ET DISPOSITIFS D'ANALYSE ET DE SYNTHÈSE AUDIO



(57) Abstract: The invention concerns a method which consists, for the analysis of an audio signal (x), in weighting samples of each frame of N samples through an analysis window of the Hamming, Hanning, Kaiser type or the like; calculating the audio signal spectrum by transforming each frame of weighted samples in the frequency domain; and in processing the spectrum of the audio signal to deliver synthesis parameters for a signal derived from an analysed audio signal. The successive frames for which complete sets of synthesis parameters are supplied advantageously have mutual overlaps of less than N/2 samples. During synthesis, the method consists in recovering the frames with an optional interpolation, and in reforming the signal by a frame overlapping sum, after having carried out an appropriate weighting of the samples.

[continued on next page]

**THIS PAGE BLANK (USPTO)**



Eurasian Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM)  
European Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB,  
GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

For an explanation of the two-letter codes and the other abbreviations, reference is made to the explanations ("Guidance Notes on Codes and Abbreviations") at the beginning of each regular edition of the PCT Gazette.

**Published:**

- With the International Search Report.

---

(57) Abrégé: Dans l'analyse du signal audio (x), on pondère les échantillons de chaque trame de N échantillons par une fenêtre d'analyse de type Hamming, Hanning, Kaiser ou analogue, on calcule un spectre du signal audio en transformant chaque trame d'échantillons pondérés dans le domaine fréquentiel, et on traite le spectre du signal audio pour délivrer des paramètres de synthèse d'un signal dérivé du signal audio analysé. Les trames successives pour lesquelles on délivre des jeux complets de paramètres de synthèse présentent avantageusement des recouvrements mutuels de moins de N/2 échantillons. A la synthèse, on récupère des trames avec une interpolation éventuelle, et on reforme le signal par une somme à recouvrement des trames, après avoir procédé à une pondération appropriée des échantillons.

PAGE BLANK (USPTO)

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
11 janvier 2001 (11.01.2001)

PCT

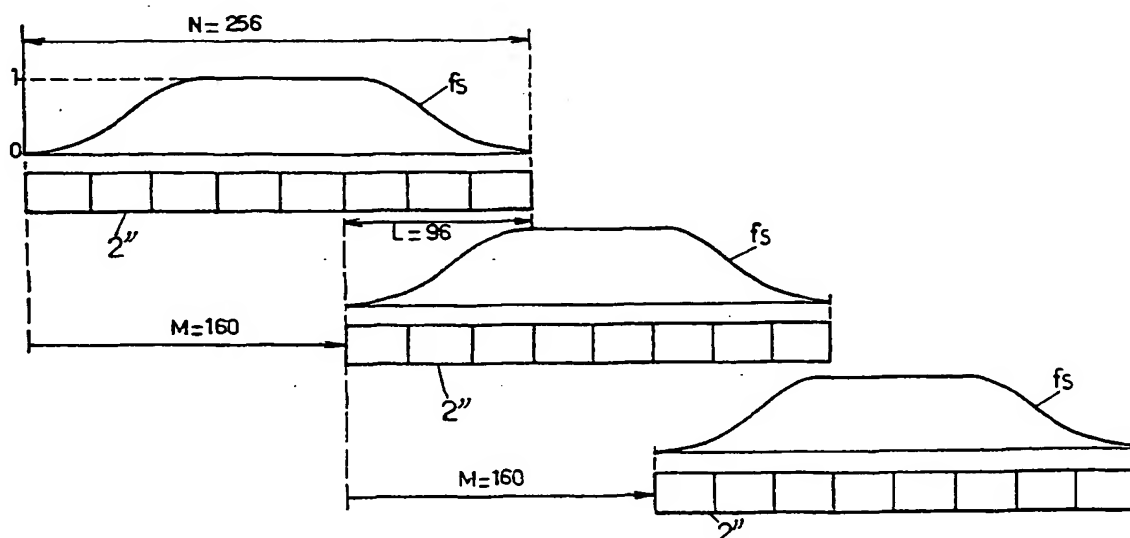
(10) Numéro de publication internationale  
**WO 01/03116 A1**

- (51) Classification internationale des brevets<sup>7</sup>: G10L 19/02
- (21) Numéro de la demande internationale:  
PCT/FR00/01904
- (22) Date de dépôt international: 4 juillet 2000 (04.07.2000)
- (25) Langue de dépôt: français
- (26) Langue de publication: français
- (30) Données relatives à la priorité:  
99/08638 5 juillet 1999 (05.07.1999) FR
- (71) Déposant (pour tous les États désignés sauf US): MA-  
TRA NORTEL COMMUNICATIONS (FR/FR); 50, rue  
du Président Sadate, F-29100 Quimper (FR).
- (72) Inventeurs; et
- (75) Inventeurs/Déposants (pour US seulement): CAPMAN,  
François (FR/FR); 47, rue des Etats Généraux, F-78000  
Versailles (FR). MURGIA, Carlo (IT/US); 458 Saint Vin-  
cent, Irvine, CA 92618 (US).
- (74) Mandataires: LOISEL, Bertrand etc.; Cabinet Plasser-  
aud, 84, rue d'Amsterdam, F-75440 Paris Cedex 09 (FR).
- (81) États désignés (national): AE, AG, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE,  
DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,  
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,  
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,  
NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,  
TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) États désignés (régional): brevet ARIPO (GH, GM, KE,  
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien

[Suite sur la page suivante]

(54) Title: METHODS AND DEVICE FOR AUDIO ANALYSIS AND SYNTHESIS

(54) Titre: PROCEDES ET DISPOSITIFS D'ANALYSE ET DE SYNTHESE AUDIO



(57) Abstract: The invention concerns a method which consists, for the analysis of an audio signal (x), in weighting samples of each frame of N samples through an analysis window of the Hamming, Hanning, Kaiser type or the like; calculating the audio signal spectrum by transforming each frame of weighted samples in the frequency domain; and in processing the spectrum of the audio signal to deliver synthesis parameters for a signal derived from an analysed audio signal. The successive frames for which complete sets of synthesis parameters are supplied advantageously have mutual overlaps of less than N/2 samples. During synthesis, the method consists in recovering the frames with an optional interpolation, and in reforming the signal by a frame overlapping sum, after having carried out an appropriate weighting of the samples.

Express Mail No. EL857974173US [Suite sur la page suivante]

WO 01/03116 A1



(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

*En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.*

**Publiée:**

— Avec rapport de recherche internationale.

---

**(57) Abrégé:** Dans l'analyse du signal audio (x), on pondère les échantillons de chaque trame de N échantillons par une fenêtre d'analyse de type Hamming, Hanning, Kaiser ou analogue, on calcule un spectre du signal audio en transformant chaque trame d'échantillons pondérés dans le domaine fréquentiel, et on traite le spectre du signal audio pour délivrer des paramètres de synthèse d'un signal dérivé du signal audio analysé. Les trames successives pour lesquelles on délivre des jeux complets de paramètres de synthèse présentent avantageusement des recouvrements mutuels de moins de N/2 échantillons. A la synthèse, on récupère des trames avec une interpolation éventuelle, et on reforme le signal par une somme à recouvrement des trames, après avoir procédé à une pondération appropriée des échantillons.

## PROCEDES ET DISPOSITIFS D'ANALYSE ET DE SYNTHESE AUDIO

La présente invention concerne l'analyse et la synthèse des signaux audio, à partir de représentations de ces signaux dans domaine spectral.

Elle s'applique en particulier, mais non exclusivement, au codage de la parole, en bande étroite ou en bande élargie, dans diverses gammes de débit de codage. Parmi les autres domaines d'application, on peut citer le débruitage par soustraction spectrale (voir EP-A-0 534 837 ou WO99/14739).

Dans les procédés d'analyse en question, le spectre du signal est obtenu en transformant des trames successives vers le domaine fréquentiel. La transformation employée est le plus souvent la transformée de Fourier rapide (TFR) ; mais d'autres transformées connues sont utilisables. Dans le cas fréquent d'un échantillonnage du signal à 8 kHz, le nombre N d'échantillons par trames est typiquement de l'ordre de 100 à 500, ce qui représente des trames de quelques dizaines de millisecondes. Pour bénéficier de la résolution maximale en fréquence, la TFR est effectuée sur 2N points, N échantillons à zéro étant ajoutés aux N échantillons de la trame.

Le spectre obtenu par transformée de Fourier de la trame de signal est la convolution du spectre réel du signal par la transformée de Fourier de la fenêtre d'analyse du signal. Cette fenêtre d'analyse, qui pondère les échantillons de chaque trame, est nécessaire à la prise en compte de la durée finie de la trame. Si la trame de signal est directement soumise à la TFR, c'est-à-dire si on utilise une fenêtre d'analyse rectangulaire, le spectre obtenu est perturbé par les pics secondaires de la TFR de la fenêtre d'analyse. Pour limiter cet inconvénient, qui est particulièrement sensible lorsque des paramètres représentant le signal ou le bruit doivent être extraits des spectres, on a recours à des fenêtres ayant de meilleures propriétés spectrales, c'est-à-dire des fonctions de pondération dont le support est limité à N échantillons et dont la transformée de Fourier a son énergie concentrée dans un pic étroit avec une forte atténuation des pics secondaires. Les plus courantes de ces fenêtres sont les fenêtres de Hamming, de Hanning et de Kaiser.

Dans la méthode d'analyse et de synthèse dite OLA (« Overlap-And-Add »), les trames successives présentent des recouvrements mutuels de 50 % (N/2 échantillons). Comme les fenêtres d'analyse couramment utilisées vérifient la propriété  $f_A(i+N/2) + f_A(i) = 1$ , la synthèse peut être effectuée simplement en effectuant la somme à recouvrement des trames de N

échantillons successivement calculées par transformée de Fourier inverse des spectres.

Dans le but d'affiner la représentation spectrale, certaines méthodes dites WOLA (« Weighted OLA ») utilisent à l'analyse des trames dont les recouvrements mutuels sont de plus de 50 %. A la synthèse, il est nécessaire de repondérer les échantillons des trames avant de les sommer. Ces méthodes augmentent la complexité de l'analyse et de la synthèse. Dans des applications de codage, elles augmentent également le débit de transmission requis.

Un but de l'invention est de proposer un schéma d'analyse et synthèse de signaux audio qui permette de limiter la cadence des trames d'analyse, tout en utilisant des fenêtres d'analyse ayant de bonnes propriétés spectrales.

L'invention propose un procédé d'analyse d'un signal audio traité par trames successives de  $N$  échantillons, dans lequel on pondère les échantillons de chaque trame par une fenêtre d'analyse de type Hamming, Hanning, Kaiser ou analogue, on calcule un spectre du signal audio en transformant chaque trame d'échantillons pondérés dans le domaine fréquentiel, et on traite le spectre du signal audio pour délivrer des paramètres de synthèse d'un signal dérivé du signal audio analysé. Suivant l'invention, les trames successives comportent une alternance de trames pour lesquelles on délivre des jeux complets de paramètres de synthèse, qui présentent des recouvrements mutuels de moins de  $N/2$  échantillons, soit moins de 50 %, et de trames pour lesquelles on délivre des jeux incomplets de paramètres de synthèse.

Les trames pour lesquelles ne sont pas délivrés des jeux complets de paramètres de synthèse peuvent ne faire l'objet d'aucune analyse spectrale. En variante, une analyse peut néanmoins être effectuée pour ces trames, afin de délivrer des jeux incomplets de paramètres de synthèse incluant des données représentant une erreur d'interpolation d'au moins un des paramètres de synthèse et/ou des données représentant un filtre d'interpolation d'au moins un des paramètres de synthèse.

Dans un premier domaine d'application du procédé, le traitement du spectre du signal audio comporte une extraction de paramètres de codage en vue de la transmission et/ou du stockage du signal audio codé. Dans un second domaine d'application du procédé, le traitement du spectre du signal audio comporte un débruitage par soustraction spectrale. D'autres domaines d'application peuvent encore être envisagés parmi les traitements audio.

Un second aspect de l'invention se rapporte à un procédé de synthèse

d'un signal audio, dans lequel on obtient des estimations spectrales successives correspondant respectivement à des trames de N échantillons du signal audio pondérés par une fenêtre d'analyse, les trames successives présentant des recouvrements mutuels de L échantillons, on évalue chaque

5 trame du signal audio en transformant les estimations spectrales dans le domaine temporel, et on combine les trames évaluées pour former le signal synthétisé. Selon ce procédé, on modifie chaque trame évaluée en lui appliquant un traitement correspondant à une division par ladite fenêtre d'analyse et à une multiplication par une fenêtre de synthèse, et on forme le

10 signal synthétisé comme une somme à recouvrement des trames modifiées. Le nombre L étant plus petit que  $N/2$  et les échantillons d'une trame ayant des rangs i numérotés de 0 à  $N-1$ , la fenêtre de synthèse  $f_S(i)$  vérifie  $f_S(N-L+i) + f_S(i) = A$  pour  $0 \leq i < L$ , et est égale à A pour  $L \leq i < N-L$ , A étant une constante positive.

15 Dans une variante du procédé de synthèse selon l'invention, on évalue un ensemble de trames recouvrantes successives de N échantillons du signal audio pondérés par une fenêtre d'analyse, en transformant dans le domaine temporel des estimations spectrales correspondant respectivement auxdites trames, et on combine les trames évaluées pour former le signal synthétisé.

20 Pour un sous-ensemble des trames évaluées, les estimations spectrales sont obtenues en traitant des paramètres de synthèse respectivement associés aux trames dudit sous-ensemble tandis que, pour les trames ne faisant pas partie du sous-ensemble, les estimations spectrales sont obtenues avec une interpolation d'une partie au moins des paramètres de synthèse. Les trames

25 successives dudit sous-ensemble présentent des décalages temporels mutuels de M échantillons, le nombre M étant plus grand que  $N/2$ , tandis que les trames successives dudit ensemble présentent des décalages temporels mutuels de  $M/p$  échantillons, p étant un entier plus grand que 1. On modifie chaque trame évaluée en lui appliquant un traitement correspondant à une division par ladite

30 fenêtre d'analyse et à une multiplication par une fenêtre de synthèse, et on forme le signal synthétisé comme une somme à recouvrement des trames modifiées. Les échantillons d'une trame ayant des rangs i numérotés de 0 à  $N-1$ , la fenêtre de synthèse  $f'_S(i)$  a un support limité aux rangs i allant de  $N/2 - M/p$  à  $N/2 + M/p$  et vérifie  $f'_S(i) + f'_S(i + M/p) = A$  pour  $N/2 - M/p \leq i < N/2$ ,

A étant une constante positive.

L'invention propose également des dispositifs de traitement audio comprenant des moyens pour la mise en œuvre des procédés d'analyse et de synthèse ci-dessus.

- 5 D'autres particularités et avantages de la présente invention apparaîtront dans la description ci-après d'exemples de réalisation non limitatifs, en référence aux dessins annexés, dans lesquels :
- la figure 1 est un schéma synoptique d'un codeur audio selon l'invention ;
  - les figures 2 et 3 sont des diagrammes illustrant la formation des trames  
10 de signal audio dans le codeur de la figure 1 ;
  - les figures 4 et 5 sont des graphiques montrant un exemple de spectre du signal audio et illustrant l'extraction des enveloppes supérieure et inférieure de ce spectre ;
  - la figure 6 est un schéma synoptique d'un exemple de moyens de  
15 quantification utilisables dans le codeur de la figure 1 ;
  - la figure 7 est un schéma synoptique de moyens utilisables pour extraire des paramètres se rapportant à la phase de la composante non-harmonique dans une variante du codeur de la figure 1 ;
  - la figure 8 est un schéma synoptique d'un décodeur audio correspondant  
20 au codeur de la figure 1 ;
  - la figure 9 est un organigramme d'un exemple de procédure de lissage de coefficients spectraux et d'extraction de phases minimales mise en œuvre dans le décodeur de la figure 8 ;
  - la figure 10 est un schéma synoptique de modules d'analyse et de  
25 mixage spectral de composantes harmonique et non-harmonique du signal audio ;
  - les figures 11 à 13 sont des graphiques montrant des exemples de fonctions non-linéaires utilisables dans le module d'analyse de la figure 10 ;
  - les figures 14 et 15 sont des diagrammes illustrant une façon de procéder  
30 à la synthèse temporelle des trames de signal dans le décodeur de la figure 8 ;
  - les figures 16 et 17 sont des graphiques montrant des fonctions de fenêtrage utilisables dans la synthèse des trames suivant les figures 14  
35 et 15 ;



- les figures 18 et 19 sont des schémas synoptiques de moyens d'interpolation utilisables dans une variante de réalisation du codeur et du décodeur ;
- la figure 20 est un schéma synoptique de moyens d'interpolation utilisables dans une autre variante de réalisation du codeur ; et
- les figures 21 et 22 sont des diagrammes illustrant une autre façon de procéder à la synthèse temporelle des trames de signal dans le décodeur de la figure 8, à l'aide d'une interpolation de paramètres.

Le codeur et le décodeur décrits ci-après sont des circuits numériques qui peuvent, comme il est usuel dans le domaine du traitement des signaux audio, être réalisés par programmation d'un processeur de signal numérique (DSP) ou d'un circuit intégré d'application spécifique (ASIC).

Le codeur audio représenté sur la figure 1 traite un signal audio d'entrée  $x$  qui, dans l'exemple non-limitatif considéré ci-après, est un signal de parole. Le signal  $x$  est disponible sous forme numérique, par exemple à une fréquence d'échantillonnage  $F_e$  de 8 kHz. Il est par exemple délivré par un convertisseur analogique-numérique traitant le signal de sortie amplifié d'un microphone. Le signal d'entrée  $x$  peut également être formé à partir d'une autre version, analogique ou numérique, codée ou non, du signal de parole.

Le codeur comprend un module 1 qui forme des trames successives de signal audio pour les différents traitements effectués, et un multiplexeur de sortie 6 qui délivre un flux de sortie  $\Phi$  contenant pour chaque trame des jeux de paramètres de quantification à partir desquels un décodeur sera capable de synthétiser une version décodée du signal audio.

La structure des trames est illustrée par les figures 2 et 3. Chaque trame 2 est composée d'un nombre  $N$  d'échantillons consécutifs du signal audio  $x$ . Les trames successives présentent des décalages temporels mutuels correspondant à  $M$  échantillons, de sorte que leur recouvrement est de  $L = N - M$  échantillons du signal. Dans l'exemple considéré, où  $N = 256$ ,  $M = 160$  et  $L = 96$ , la durée des trames 2 est de  $N/F_e = 32$  ms, et une trame est formée toutes les  $M/F_e = 20$  ms.

De façon classique, le module 1 multiplie les échantillons de chaque trame 2 par une fonction de fenêtrage  $f_A$ , de préférence choisie pour ses bonnes propriétés spectrales. Les échantillons  $x(i)$  de la trame étant numérotés de  $i = 0$  à  $i = N - 1$ , la fenêtre d'analyse  $f_A(i)$  peut ainsi être une fenêtre de

Hamming, d'expression :

$$f_A(i) = 0,54 + 0,46 \cdot \cos\left(2\pi \frac{i - (N-1)/2}{N}\right) \quad (1)$$

ou une fenêtre de Hanning d'expression :

$$f_A(i) = \frac{1}{2} \left( 1 + \cos\left(2\pi \frac{i - (N-1)/2}{N}\right) \right) \quad (2)$$

5 ou encore une fenêtre de Kaiser, d'expression :

$$f_A(i) = \frac{I_0\left(\alpha \sqrt{1 - \left(\frac{i - (N-1)/2}{N}\right)^2}\right)}{I_0(\alpha)} \quad (3)$$

où  $\alpha$  est un coefficient par exemple égal à 6, et  $I_0(\cdot)$  désigne la fonction de Bessel d'indice 0.

Le codeur de la figure 1 procède à une analyse du signal audio dans le  
 10 domaine spectral. Il comporte un module 3 qui calcule la transformée de  
 Fourier rapide (TFR) de chaque trame de signal. La trame de signal est mise  
 en forme avant d'être soumise au module de TFR 3 : le module 1 lui adjoint  
 $N = 256$  échantillons à zéro afin d'obtenir la résolution maximale de la  
 transformée de Fourier, et il effectue d'autre part une permutation circulaire des  
 15  $2N = 512$  échantillons afin de compenser les effets de phase résultant de la  
 fenêtre d'analyse. Cette modification de la trame est illustrée par la figure 3. La  
 trame dont on calcule la transformée de Fourier rapide sur  $2N = 512$  points  
 commence par les  $N/2 = 128$  derniers échantillons pondérés de la trame, suivis  
 par les  $N = 256$  échantillons à zéro, et se termine par les  $N/2 = 128$  premiers  
 20 échantillons pondérés de la trame.

Le module de TFR 3 obtient le spectre du signal pour chaque trame,  
 dont le module et la phase sont respectivement notés  $|X|$  et  $\varphi_X$ , ou  $|X(i)|$  et  
 $\varphi_X(i)$  pour les index de fréquence  $i = 0$  à  $i = 2N-1$  (grâce à la symétrie de la  
 transformée de Fourier et des trames, on peut se limiter aux valeurs pour  
 25  $0 \leq i < N$ ).

Un détecteur de fréquence fondamentale 4 estime pour chaque trame  
 de signal une valeur de la fréquence fondamentale  $F_0$ . Le détecteur 4 peut  
 appliquer toute méthode connue d'analyse du signal de parole de la trame pour  
 estimer la fréquence fondamentale  $F_0$ , par exemple une méthode basée sur la

fonction d'autocorrélation ou la fonction AMDF, éventuellement précédée d'un module de blanchiment par prédiction linéaire. L'estimation peut également être effectuée dans le domaine spectral ou dans le domaine cepstral. Une autre possibilité est d'évaluer les intervalles de temps entre les ruptures consécutives du signal de parole attribuables à des fermetures de la glotte du locuteur intervenant pendant la durée de la trame. Des méthodes bien connues utilisables pour détecter de telles micro-ruptures sont décrites dans les articles suivants : M. Basseville et al., « Sequential detection of abrupt changes in spectral characteristics of digital signals » (IEEE Trans. on Information Theory, 1983, Vol. IT-29, n° 5, pages 708-723) ; R. Andre-Obrecht, « A new statistical approach for the automatic segmentation of continuous speech signals » (IEEE Trans. on Acous., Speech and Sig. Proc., Vol. 36, N°1, janvier 1988) ; et C. MURGIA et al., « An algorithm for the estimation of glottal closure instants using the sequential detection of abrupt changes in speech signals » (Signal Processing VII, 1994, pages 1685-1688).

La fréquence fondamentale estimée  $F_0$  fait l'objet d'une quantification, par exemple scalaire, par un module 5, qui fournit au multiplexeur de sortie 6 un index iF de quantification de la fréquence fondamentale pour chaque trame du signal.

Le codeur utilise des modélisations paramétriques cepstrales pour représenter une enveloppe supérieure et une enveloppe inférieure du spectre du signal audio. La première étape de la transformation cepstrale consiste à appliquer au module du spectre du signal une fonction de compression spectrale, qui peut être une fonction logarithmique ou en racine. Le module 8 du codeur opère ainsi, pour chaque valeur  $X(i)$  du spectre du signal ( $0 \leq i < N$ ), la transformation suivante :

$$LX(i) = \text{Log}(|X(i)|) \quad (4)$$

dans le cas d'une compression logarithmique ou

$$LX(i) = |X(i)|^\gamma \quad (5)$$

dans le cas d'une compression en racine,  $\gamma$  étant un exposant compris entre 0 et 1.

Le spectre comprimé LX du signal audio est traité par un module 9 qui extrait des amplitudes spectrales associées aux harmoniques du signal correspondant aux multiples de la fréquence fondamentale estimée  $F_0$ . Ces

amplitudes sont ensuite interpolées par un module 10 afin d'obtenir une enveloppe supérieure comprimée notée LX<sub>sup</sub>.

Il est à noter que la compression spectrale pourrait de façon équivalente être effectuée après la détermination des amplitudes associées aux harmoniques. Elle pourrait également être effectuée après l'interpolation, ce qui ne ferait que modifier la forme des fonctions d'interpolation.

Le module 9 d'extraction des maxima tient compte de l'éventuelle variation de la fréquence fondamentale sur la trame d'analyse, des erreurs que peut commettre le détecteur 4, ainsi que des imprécisions liées au caractère discret de l'échantillonnage en fréquence. Pour cela, la recherche des amplitudes des pics spectraux ne consiste pas simplement à prendre les valeurs LX(i) correspondant aux index i tels que  $i.F_e/2N$  soit la fréquence la plus proche d'une harmonique de fréquence  $k.F_0$  ( $k \geq 1$ ). L'amplitude spectrale retenue pour une harmonique d'ordre k est un maximum local du module du spectre au voisinage de la fréquence  $k.F_0$  (cette amplitude est obtenue directement sous forme comprimée lorsque la compression spectrale 8 est effectuée avant l'extraction des maxima 9).

Les figures 4 et 5 montrent un exemple de forme du spectre comprimé LX, où on voit que les amplitudes maximales des pics harmoniques ne coïncident pas nécessairement avec les amplitudes correspondant aux multiples entiers de la fréquence fondamentale estimée  $F_0$ . Les flancs des pics étant assez raides, une petite erreur de positionnement de la fréquence fondamentale  $F_0$ , amplifiée par l'indice d'harmonique k, peut distordre fortement l'enveloppe supérieure estimée du spectre et provoquer une mauvaise modélisation de la structure formantique du signal. Par exemple, prendre directement l'amplitude spectrale pour la fréquence  $3.F_0$  dans le cas des figures 4 et 5 produirait une erreur importante dans l'extraction de l'enveloppe supérieure au voisinage de l'harmonique d'ordre  $k = 3$ , alors qu'il s'agit d'une zone énergétiquement importante dans l'exemple dessiné. En effectuant l'interpolation à partir du véritable maximum, on évite ce genre d'erreur d'estimation de l'enveloppe supérieure.

Dans l'exemple représenté sur la figure 4, l'interpolation est effectuée entre des points dont l'abscisse est la fréquence correspondant au maximum de l'amplitude d'un pic spectral, et dont l'ordonnée est ce maximum, avant ou

après compression.

L'interpolation effectuée pour calculer l'enveloppe supérieure LX<sub>sup</sub> est une simple interpolation linéaire. Bien entendu une autre forme d'interpolation pourrait être utilisée (par exemple polynomiale ou spline).

5 Dans la variante préférée représentée sur la figure 5, l'interpolation est effectuée entre des points dont l'abscisse est une fréquence  $k.F_0$  multiple de la fréquence fondamentale (en fait la fréquence la plus proche dans le spectre discret) et dont l'ordonnée est l'amplitude maximale, avant ou après compression, du spectre au voisinage de cette fréquence multiple.

10 En comparant les figures 4 et 5, on peut voir que le mode d'extraction selon la figure 5, qui repositionne les pics sur les fréquences harmoniques, conduit à une meilleure précision sur l'amplitude des pics que le décodeur attribuera aux fréquences multiples de la fréquence fondamentale. Il peut se produire un léger déplacement en fréquence de la position de ces pics, ce qui  
15 n'est pas perceptuellement très important et n'est d'ailleurs pas évité non plus dans le cas de la figure 4. Dans le cas de la figure 4, les points d'ancrage pour l'interpolation sont confondus avec les sommets des pics harmoniques. Dans le cas de la figure 5, on impose que ces points d'ancrage se trouvent précisément aux fréquences multiples de la fréquence fondamentale, leurs amplitudes  
20 correspondant à celles des pics.

L'intervalle de recherche du maximum d'amplitude associé à une harmonique de rang  $k$  est centré sur l'index  $i$  de la fréquence de la TFR la plus proche de  $k.F_0$ , c'est-à-dire  $i = \left\lfloor 2Nk \frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$ , où  $[a]$  désigne l'entier égal ou  
25 immédiatement inférieur au nombre  $a$ . La largeur de cet intervalle de recherche dépend de la fréquence d'échantillonnage  $F_e$ , de la taille  $2N$  de la TFR et de la gamme de variation possible de la fréquence fondamentale. Cette largeur est typiquement de l'ordre d'une dizaine de fréquences avec les exemples de valeurs précédemment considérés. On peut la rendre réglable en fonction de la valeur  $F_0$  de la fréquence fondamentale et du numéro  $k$  de l'harmonique.

30 Afin d'améliorer la résolution dans les basses fréquences et donc de représenter plus fidèlement les amplitudes des harmoniques dans cette zone, une distorsion non-linéaire de l'échelle des fréquences est opérée sur l'enveloppe supérieure comprimée par un module 12 avant que le module 13 effectue la transformée de Fourier rapide inverse (TFRI) fournissant les

coefficients cepstraux  $cx\_sup$ .

La distorsion non-linéaire permet de minimiser plus efficacement l'erreur de modélisation. Elle est par exemple effectuée selon une échelle de fréquences de type Mel ou Bark. Cette distorsion peut éventuellement dépendre de la fréquence fondamentale estimée  $F_0$ . La figure 1 illustre le cas de l'échelle Mel. La relation entre les fréquences  $F$  du spectre linéaire, exprimées en hertz, et les fréquences  $F'$  de l'échelle Mel est la suivante :

$$F' = \frac{1000}{\log_{10}(2)} \times \log_{10} \left( 1 + \frac{F}{1000} \right) \quad (6)$$

Afin de limiter le débit de transmission, une troncature des coefficients cepstraux  $cx\_sup$  est effectuée. Le module de TFR1 13 a besoin de calculer seulement un vecteur cepstral de NCS coefficients cepstraux d'ordres 0 à NCS-1. A titre d'exemple, NCS peut être égal à 16.

Un post-filtrage dans le domaine cepstral, appelé post-liffrage, est appliqué par un module 15 à l'enveloppe supérieure comprimée  $LX\_sup$ . Ce post-liffrage correspond à une manipulation des coefficients cepstraux  $cx\_sup$  délivrés par le module de TFR1 13, qui correspond approximativement à un post-filtrage de la partie harmonique du signal par une fonction de transfert ayant la forme classique :

$$H(z) = \left( 1 - \mu z^{-1} \right) \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (7)$$

où  $A(z)$  est la fonction de transfert d'un filtre de prédiction linéaire du signal audio,  $\gamma_1$  et  $\gamma_2$  sont des coefficients compris entre 0 et 1, et  $\mu$  est un coefficient de préaccentuation éventuellement nul. La relation entre le coefficient post-liffré d'ordre  $i$ , noté  $c_p(i)$ , et le coefficient cepstral correspondant  $c(i) = cx\_sup(i)$  délivré par le module 13 est alors :

$$\begin{aligned} c_p(0) &= c(0) \\ c_p(i) &= \left( 1 + \gamma_2^i - \gamma_1^i \right) c(i) - \frac{\mu^i}{i} \quad \text{pour } i > 0 \end{aligned} \quad (8)$$

Le coefficient de préaccentuation optionnel  $\mu$  peut être contrôlé en posant comme contrainte de préserver la valeur du coefficient cepstral  $cx\_sup(1)$  relatif à la pente. En effet, la valeur  $c(1) = cx\_sup(1)$  d'un bruit blanc filtré par le filtre de préaccentuation correspond au coefficient de préaccentuation. On peut ainsi choisir ce dernier de la façon suivante :

$$\mu = (\gamma_2 - \gamma_1) \cdot c(1).$$

Après le post-lifre 15, un module de normalisation 16 modifie encore les coefficients cepstraux en imposant la contrainte de modélisation exacte d'un point du spectre initial, qui est de préférence le point le plus énergétique parmi les maxima spectraux extraits par le module 9. En pratique, cette  
 5 normalisation modifie seulement la valeur du coefficient  $c_p(0)$ .

Le module de normalisation 16 fonctionne de la façon suivante : il recalcule une valeur du spectre synthétisé à la fréquence du maximum indiqué par le module 9, par transformée de Fourier des coefficients cepstraux tronqués et post-lifrés, en tenant compte de la distorsion non-linéaire de l'axe  
 10 des fréquences ; il détermine un gain de normalisation  $g_N$  par la différence logarithmique entre la valeur du maximum fournie par le module 9 et cette valeur recalculée ; et il ajoute le gain  $g_N$  au coefficient cepstral post-lifré  $c_p(0)$ . Cette normalisation peut être vue comme faisant partie du post-lifrage.

Les coefficients cepstraux post-lifrés et normalisés font l'objet d'une  
 15 quantification par un module 18 qui transmet des index de quantification correspondants icxs au multiplexeur de sortie 6 du codeur.

Le module 18 peut fonctionner par quantification vectorielle à partir de vecteurs cepstraux formés de coefficients post-lifrés et normalisés, notés ici  $cx[n]$  pour la trame de signal de rang  $n$ . A titre d'exemple, le vecteur cepstral  
 20  $cx[n]$  de NCS = 16 coefficients cepstraux  $cx[n,0]$ ,  $cx[n,1]$ , ...,  $cx[n,NCS-1]$  est distribué en quatre sous-vecteurs cepstraux contenant chacun quatre coefficients d'ordres consécutifs. Le vecteur cepstral  $cx[n]$  peut être traité par les moyens représentés sur la figure 6, faisant partie du module de quantification 18. Ces moyens mettent en œuvre, pour chaque composante  
 25  $cx[n,i]$ , un prédicteur de la forme :

$$cx_p[n,i] = (1 - \alpha(i)).rcx[n,i] + \alpha(i).rcx[n-1,i] \quad (9)$$

où  $rcx[n]$  désigne un vecteur résiduel de prédiction pour la trame de rang  $n$  dont les composantes sont respectivement notées  $rcx[n,0]$ ,  $rcx[n,1]$ , ...,  $rcx[n,NCS-1]$ , et  $\alpha(i)$  désigne un coefficient de prédiction choisi pour être  
 30 représentatif d'une corrélation inter-trame supposée. Après quantification des résidus, ce vecteur résiduel est défini par :

$$rcx[n,i] = \frac{cx[n,i] - \alpha(i).rcx\_q[n-1,i]}{2 - \alpha(i)} \quad (10)$$

où  $rcx\_q[n-1]$  désigne le vecteur résiduel quantifié pour la trame de rang  $n-1$ ,

dont les composantes sont respectivement notées  $rcx\_q[n,0]$ ,  $rcx\_q[n,1]$ , ...,  $rcx\_q[n,NCS-1]$ .

Le numérateur de la relation (10) est obtenu par un soustracteur 20, dont les composantes du vecteur de sortie sont divisées par les quantités  $2-\alpha(i)$  en 21. Aux fins de la quantification, le vecteur résiduel  $rcx[n]$  est subdivisé en quatre sous-vecteurs, correspondant à la subdivision en quatre sous-vecteurs cepstraux. Sur la base d'un dictionnaire obtenu par apprentissage préalable, l'unité 22 procède à la quantification vectorielle de chaque sous-vecteur du vecteur résiduel  $rcx[n]$ . Cette quantification peut consister, pour chaque sous-vecteur  $srcx[n]$ , à sélectionner dans le dictionnaire le sous-vecteur quantifié  $srcx\_q[n]$  qui minimise l'erreur quadratique  $\|srcx[n] - srcx\_q[n]\|^2$ . L'ensemble  $icxs$  des index de quantification  $icx$ , correspondant aux adresses dans le ou les dictionnaires des sous-vecteurs résiduels quantifiés  $srcx\_q[n]$ , est fourni au multiplexeur de sortie 6.

L'unité 22 délivre également les valeurs des sous-vecteurs résiduels quantifiés, qui forment le vecteur  $rcx\_q[n]$ . Celui-ci est retardé d'une trame en 23, et ses composantes sont multipliées par les coefficients  $\alpha(i)$  en 24 pour fournir le vecteur à l'entrée négative du soustracteur 20. Ce dernier vecteur est d'autre part fourni à un additionneur 25, dont l'autre entrée reçoit un vecteur formé par les composantes du résidu quantifié  $rcx\_q[n]$  respectivement multipliées par les quantités  $1-\alpha(i)$  en 26. L'additionneur 25 délivre ainsi le vecteur cepstral quantifié  $cx\_q[n]$  que récupérera le décodeur.

Le coefficient de prédiction  $\alpha(i)$  peut être optimisé séparément pour chacun des coefficients cepstraux. Les dictionnaires de quantification peuvent aussi être optimisés séparément pour chacun quatre sous-vecteurs cepstraux. D'autre part, il est possible, de façon connue en soi, de normaliser les vecteurs cepstraux avant d'appliquer le schéma de prédiction/quantification, à partir de la variance des cepstres.

Il est à noter que le schéma ci-dessus de quantification des coefficients cepstraux peut n'être appliqué que pour certaines seulement des trames. Par exemple, on peut prévoir un second mode de quantification ainsi qu'un processus de sélection de celui des deux modes qui minimise un critère de moindres carrés avec les coefficients cepstraux à quantifier, et transmettre avec les index de quantification de la trame un bit indiquant lequel des deux modes a été sélectionné.



Les coefficients cepstraux quantifiés  $cx\_sup\_q = cx\_q[n]$  fournis par l'additionneur 25 sont adressés à un module 28 qui recalcule les amplitudes spectrales associées à une ou plusieurs des harmoniques de la fréquence fondamentale  $F_0$  (figure 1). Ces amplitudes spectrales sont par exemple

5 calculées sous forme comprimée, en appliquant la transformée de Fourier aux coefficients cepstraux quantifiés en tenant compte de la distorsion non-linéaire de l'échelle des fréquences utilisée dans la transformation cepstrale. Les amplitudes ainsi recalculées sont fournies à un module d'adaptation 29 qui les compare à des amplitudes de maxima déterminées par le module d'extraction

10 9.

Le module d'adaptation 29 contrôle le post-lifre 15 de façon à minimiser un écart de module entre le spectre du signal audio et les valeurs de module correspondantes calculées en 28. Cet écart de module peut être exprimé par une somme de valeurs absolues de différences d'amplitudes,

15 comprimées ou non, correspondant à une ou plusieurs des fréquences harmoniques. Cette somme peut être pondérée en fonction des amplitudes spectrales associées à ces fréquences.

De façon optimale, l'écart de module pris en compte dans l'adaptation du post-lifrage tiendrait compte de toutes les harmoniques du spectre. Cependant, afin de réduire la complexité de l'optimisation, le module 28 peut

20 ne resynthétiser les amplitudes spectrales que pour une ou plusieurs fréquences multiples de la fréquence fondamentale  $F_0$ , sélectionnées sur la base de l'importance du module du spectre en valeur absolue. Le module d'adaptation 29 peut par exemple considérer les trois pics spectraux les plus

25 intenses dans le calcul de l'écart de module à minimiser.

Dans une autre réalisation, le module d'adaptation 29 estime une courbe de masquage spectral du signal audio au moyen d'un modèle psychoacoustique, et les fréquences prises en compte dans le calcul de l'écart de module à minimiser sont sélectionnées sur la base de l'importance du

30 module du spectre relativement à la courbe de masquage (on peut par exemple prendre les trois fréquences pour lesquelles le module du spectre dépasse le plus de la courbe de masquage). Différentes méthodes classiques sont utilisables pour calculer la courbe de masquage à partir du signal audio. On peut par exemple utiliser celle développée par J.D. Johnston (« Transform

35 Coding of Audio Signals Using Perceptual Noise Criteria », IEEE Journal on

Selected Area in Communications, Vol. 6, No. 2, février 1988).

Pour réaliser l'adaptation du post-lifrage, le module 29 peut utiliser un modèle d'identification de filtre. Une méthode plus simple consiste à prédéfinir un ensemble de jeux de paramètres de post-lifrage, c'est-à-dire un ensemble  
5 de couples  $\gamma_1, \gamma_2$  dans le cas d'un post-lifrage selon les relations (8), à effectuer les opérations incombant aux modules 15, 16, 18 et 28 pour chacun de ces jeux de paramètres, et à retenir celui des jeux de paramètres qui conduit à l'écart de module minimal entre le spectre du signal et les valeurs recalculées. Les index de quantification fournis par le module 18 sont alors  
10 ceux qui se rapportent au meilleur jeu de paramètres.

Par un processus analogue à celui de l'extraction des coefficients  $cx\_sup$  représentant l'enveloppe supérieure comprimée  $LX\_sup$  du spectre du signal, le codeur détermine des coefficients  $cx\_inf$  représentant une enveloppe inférieure comprimée  $LX\_inf$ . Un module 30 extrait du spectre comprimé  $LX$   
15 des amplitudes spectrales associées à des fréquences situées dans des zones du spectre intermédiaires par rapport aux fréquences multiples de la fréquence fondamentale estimée  $F_0$ .

Dans l'exemple illustré par les figures 4 et 5, chaque amplitude associée à une fréquence située dans une zone intermédiaire entre deux  
20 harmoniques successives  $k.F_0$  et  $(k+1).F_0$  correspond simplement au module du spectre pour la fréquence  $(k+1/2).F_0$  située au milieu de l'intervalle séparant les deux harmoniques. Dans une autre réalisation, cette amplitude pourrait être une moyenne du module du spectre sur une petite plage entourant cette fréquence  $(k+1/2).F_0$ .

25 Un module 31 procède à une interpolation, par exemple linéaire, des amplitudes spectrales associées aux fréquences situées dans les zones intermédiaires pour obtenir l'enveloppe inférieure comprimée  $LX\_inf$ .

La transformation cepstrale appliquée à cette enveloppe inférieure comprimée  $LX\_inf$  est effectuée suivant une échelle de fréquences résultant  
30 d'une distorsion non-linéaire appliquée par un module 32. Le module de TFRI 33 calcule un vecteur cepstral de  $NCI$  coefficients cepstraux  $cx\_inf$  d'ordres 0 à  $NCI-1$  représentant l'enveloppe inférieure.  $NCI$  est un nombre qui peut être sensiblement plus petit que  $NCS$ , par exemple  $NCI = 4$ .

35 La transformation non-linéaire de l'échelle des fréquences pour la transformation cepstrale de l'enveloppe inférieure peut être réalisée vers une

échelle plus fine aux hautes fréquences qu'aux basses fréquences, ce qui permet avantageusement de bien modéliser les composantes non-voisées du signal aux hautes fréquences. Toutefois, pour assurer une homogénéité de représentation entre l'enveloppe supérieure et l'enveloppe inférieure, on pourra  
5 préférer adopter dans le module 32 la même échelle que dans le module 12 (Mel dans l'exemple considéré).

Les coefficients cepstraux  $cx\_inf$  représentant l'enveloppe inférieure comprimée sont quantifiés par un module 34, qui peut fonctionner de la même manière que le module 18 de quantification des coefficients cepstraux  
10 représentant l'enveloppe supérieure comprimée. Dans le cas considéré, où on se limite à  $NCI = 4$  coefficients cepstraux pour l'enveloppe inférieure, le vecteur ainsi formé est soumis à une quantification vectorielle de résidu de prédiction, effectuée par des moyens identiques à ceux représentés sur la figure 6 mais sans subdivision en sous-vecteurs. L'index de quantification  $icx = icxi$   
15 déterminé par le quantificateur vectoriel 22 pour chaque trame relativement aux coefficients  $cx\_inf$  est fourni au multiplexeur de sortie 6 du codeur.

Le codeur représenté sur la figure 1 ne comporte aucun dispositif particulier pour coder les phases du spectre aux harmoniques du signal audio.

En revanche, il comporte des moyens 36-40 pour coder une  
20 information temporelle liée à la phase de la composante non-harmonique représentée par l'enveloppe inférieure.

Un module 36 de décompression spectrale et un module 37 de TFRI forment une estimation temporelle de la trame de la composante non-harmonique. Le module 36 applique une fonction de décompression réciproque  
25 de la fonction de compression appliquée par le module 8 (c'est-à-dire une exponentielle ou une fonction puissance  $1/\gamma$ ) à l'enveloppe inférieure comprimée  $LX\_inf$  produite par le module d'interpolation 31. Ceci fournit le module de la trame estimée de la composante non-harmonique, dont la phase est prise égale à celle  $\phi_X$  du spectre du signal X sur la trame. La transformée  
30 de Fourier inverse effectuée par le module 37 fournit la trame estimée de la composante non-harmonique.

Le module 38 subdivise cette trame estimée de la composante non-harmonique en plusieurs segments temporels. La trame délivrée par le module 37 se composant de  $2N = 512$  échantillons pondérés comme illustré par la  
35 figure 3, le module 38 considère seulement les  $N/2 = 128$  premiers échantillons

et les  $N/2 = 128$  derniers échantillons, et les subdivise par exemple en huit segments de 32 échantillons consécutifs représentant chacun 4 ms de signal.

Pour chaque segment, le module 38 calcule l'énergie égale à la  
5 somme des carrés des échantillons, et forme un vecteur E1 formé de  
huit composantes réelles positives égales aux huit énergies calculées. La plus  
grande de ces huit énergies, notée EM, est également déterminée pour être  
fournie, avec le vecteur E1, à un module de normalisation 39. Celui-ci divise  
10 chaque composante du vecteur E1 par EM, de sorte que le vecteur normalisé  
Emix est formé de huit composantes comprises entre 0 et 1. C'est ce vecteur  
normalisé Emix, ou vecteur de pondération, qui est soumis à la quantification  
par le module 40. Celui-ci peut opérer une quantification vectorielle avec un  
dictionnaire déterminé lors d'un apprentissage préalable. L'index de  
quantification iEm est fourni par le module 40 au multiplexeur de sortie 6 du  
15 codeur.

La figure 7 montre une variante de réalisation des moyens employés  
par le codeur de la figure 1 pour déterminer le vecteur Emix de pondération  
énergétique de la trame de la composante non-harmonique. Les modules 36,  
37 de décompression spectrale et de TFR1 fonctionnent comme ceux qui  
20 portent les mêmes références sur la figure 1. Un module de sélection 42 est  
ajouté pour déterminer la valeur du module du spectre soumis à la transformée  
de Fourier inverse 37. Sur la base de la fréquence fondamentale estimée  $F_0$ , le  
module 42 identifie des régions harmoniques et des régions non-harmoniques  
du spectre du signal audio. Par exemple, une fréquence sera considérée  
25 comme appartenant à une région harmonique si elle se trouve dans un  
intervalle de fréquences centré sur une harmonique  $k.F_0$  et de largeur  
correspondant à une largeur de raie spectrale synthétisée, et à une région non-  
harmonique sinon. Dans les régions non-harmoniques, le signal complexe  
soumis à la TFR1 37 est égal à la valeur du spectre, c'est-à-dire que son  
30 module et sa phase correspondent aux valeurs  $|X|$  et  $\phi_X$  fournies par le module  
de TFR 3. Dans les régions harmoniques, ce signal complexe a la même  
phase  $\phi_X$  que le spectre et un module donné par l'enveloppe inférieure après  
décompression spectrale 36. Cette façon de procéder selon la figure 7 procure  
une modélisation plus précise des régions non-harmoniques.

35 Le décodeur représenté sur la figure 8 comprend un démultiplexeur

d'entrée 45 qui extrait du flux binaire  $\Phi$ , issu d'un codeur selon la figure 1, les index  $iF$ ,  $icxs$ ,  $icxi$ ,  $iEm$  de quantification de la fréquence fondamentale  $F_0$ , des coefficients cepstraux représentant l'enveloppe supérieure comprimée, des coefficients représentant l'enveloppe inférieure comprimée, et du vecteur de pondération  $Emix$ , et les distribue respectivement à des modules 46, 47, 48 et 49. Ces modules 46-49 comportent des dictionnaires de quantification semblables à ceux des modules 5, 18, 34 et 40 de la figure 1, afin de restituer les valeurs des paramètres quantifiés. Les modules 47 et 48 ont des dictionnaires pour former les résidus de prédiction quantifiés  $rcx\_q[n]$ , et ils en déduisent les vecteurs cepstraux quantifiés  $cx\_q[n]$  avec des éléments identiques aux éléments 23-26 de la figure 6. Ces vecteurs cepstraux quantifiés  $cx\_q[n]$  fournissent les coefficients cepstraux  $cx\_sup\_q$  et  $cx\_inf\_q$  traités par le décodeur.

Un module 51 calcule la transformée de Fourier rapide des coefficients cepstraux  $cx\_sup$  pour chaque trame de signal. L'échelle des fréquences du spectre comprimé qui en résulte est modifiée non-linéairement par un module 52 appliquant la transformation non-linéaire réciproque de celle du module 12 de la figure 1, et qui fournit l'estimation  $LX\_sup$  de l'enveloppe supérieure comprimée. Une décompression spectrale de  $LX\_sup$ , opérée par un module 53, fournit l'enveloppe supérieure  $X\_sup$  comportant les valeurs estimées du module du spectre aux fréquences multiples de la fréquence fondamentale  $F_0$ . Le module 54 synthétise l'estimation spectrale  $X_v$  de la composante harmonique du signal audio, par une somme de raies spectrales centrées sur les fréquences multiples de la fréquence fondamentale  $F_0$  et dont les amplitudes (en module) sont celles données par l'enveloppe supérieure  $X\_sup$ .

Bien que le flux numérique d'entrée  $\Phi$  ne comporte pas d'informations spécifiques sur la phase du spectre du signal aux harmoniques de la fréquence fondamentale, le décodeur de la figure 8 est capable d'extraire de l'information sur cette phase à partir des coefficients cepstraux  $cx\_sup\_q$  représentant l'enveloppe supérieure comprimée. Cette information de phase est utilisée pour affecter une phase  $\phi(k)$  à chacune des raies spectrales déterminées par le module 54 dans l'estimation de la composante harmonique du signal.

En première approximation, le signal de parole peut être considéré comme étant à phase minimale. D'autre part, il est connu que l'information de phase minimale peut se déduire facilement d'une modélisation cepstrale. Cette

information de phase minimale est donc calculée pour chaque fréquence harmonique. L'hypothèse de phase minimale signifie que l'énergie du signal synthétisé est localisée au début de chaque période de la fréquence fondamentale  $F_0$ .

5 Pour être plus proche d'un signal de parole réel, on introduit un peu de dispersion au moyen d'un post-liffrage spécifique des cepstres lors de la synthèse de la phase. Avec ce post-liffrage, effectué par le module 55 de la figure 8, il est possible d'accentuer les résonances formantiques de l'enveloppe et donc de contrôler la dispersion des phases. Ce post-liffrage est par exemple  
10 de la forme (8).

Pour limiter les ruptures de phase, il est préférable de lisser les coefficients cepstraux post-liftrés, ce qui est effectué par le module 56. Le module 57 déduit des coefficients cepstraux post-liftrés et lissés la phase minimale affectée à chaque raie spectrale représentant un pic harmonique du  
15 spectre.

Les opérations effectuées par les modules 56, 57 de lissage et d'extraction de la phase minimale sont illustrées par l'organigramme de la figure 9. Le module 56 examine les variations des coefficients cepstraux pour appliquer un lissage moins important en présence de variations brusques qu'en  
20 présence de variations lentes. Pour cela, il effectue le lissage des coefficients cepstraux au moyen d'un facteur d'oubli  $\lambda_c$  choisi en fonction d'une comparaison entre un seuil  $d_{th}$  et une distance  $d$  entre deux jeux successifs de coefficients cepstraux post-liftrés. Le seuil  $d_{th}$  est lui-même adapté en fonction des variations des coefficients cepstraux.

25 La première étape 60 consiste à calculer la distance  $d$  entre les deux vecteurs successifs relatifs aux trames  $n-1$  et  $n$ . Ces vecteurs, notés ici  $c_{xp}[n-1]$  et  $c_{xp}[n]$ , correspondent pour chaque trame à l'ensemble des NCS coefficients cepstraux post-liftrés représentant l'enveloppe supérieure comprimée. La distance utilisée peut notamment être la distance euclidienne  
30 entre les deux vecteurs ou encore une distance quadratique.

Deux lissages sont d'abord effectués, respectivement au moyen de facteurs d'oubli  $\lambda_{min}$  et  $\lambda_{max}$ , pour déterminer une distance minimale  $d_{min}$  et une distance maximale  $d_{max}$ . Le seuil  $d_{th}$  est ensuite déterminé à l'étape 70 comme étant situé entre les distances minimale et maximale  $d_{min}$ ,  $d_{max}$  :

$d_{th} = \beta \cdot d_{max} + (1-\beta) \cdot d_{min}$ , le coefficient  $\beta$  étant par exemple égal à 0,5.

Dans l'exemple représenté, les facteurs d'oubli  $\lambda_{min}$  et  $\lambda_{max}$  sont eux-mêmes sélectionnés parmi deux valeurs distinctes, respectivement  $\lambda_{min1}$ ,  $\lambda_{min2}$  et  $\lambda_{max1}$ ,  $\lambda_{max2}$  comprises entre 0 et 1, les indices  $\lambda_{min1}$ ,  $\lambda_{max1}$  étant chacun  
 5 sensiblement plus près de 0 que les indices  $\lambda_{min2}$ ,  $\lambda_{max2}$ . Si  $d > d_{min}$  (test 61), le facteur d'oubli  $\lambda_{min}$  est égal à  $\lambda_{min1}$  (étape 62) ; sinon il est pris égal à  $\lambda_{min2}$  (étape 63). A l'étape 64, la distance minimale  $d_{min}$  est prise égale à  $\lambda_{min} \cdot d_{min} + (1-\lambda_{min}) \cdot d$ . Si  $d > d_{max}$  (test 65), le facteur d'oubli  $\lambda_{max}$  est égal à  $\lambda_{max1}$  (étape 66) ; sinon il est pris égal à  $\lambda_{max2}$  (étape 67). A l'étape 68, la  
 10 distance minimale  $d_{max}$  est prise égale à  $\lambda_{max} \cdot d_{max} + (1-\lambda_{max}) \cdot d$ .

Si la distance  $d$  entre les deux vecteurs cepstraux consécutifs est plus grande que le seuil  $d_{th}$  (test 71), on adopte pour le facteur d'oubli  $\lambda_c$  une valeur  $\lambda_{c1}$  relativement proche de 0 (étape 72). On considère dans ce cas que le  
 15 signal correspondant est de type non stationnaire, de sorte qu'il n'y a pas lieu de conserver une grande mémoire des coefficients cepstraux antérieurs. Si  $d \leq d_{th}$ , on adopte à l'étape 73 pour le facteur d'oubli  $\lambda_c$  une valeur  $\lambda_{c2}$  moins proche de 0 afin de lisser davantage les coefficients cepstraux. Le lissage est effectué à l'étape 74, où le vecteur  $cxl[n]$  de coefficients lissés pour la trame courante  $n$  est déterminé par :

$$20 \quad cxl[n] = \lambda_c \cdot cxl[n-1] + (1-\lambda_c) \cdot cxp[n] \quad (11)$$

Le module 57 calcule ensuite les phases minimales  $\varphi(k)$  associées aux harmoniques  $k \cdot F_0$ . De façon connue, la phase minimale pour une harmonique d'ordre  $k$  est donnée par :

$$\varphi(k) = -2 \cdot \sum_{m=1}^{NCS-1} cxl[n,m] \cdot \sin(2\pi m k F_0 / F_e) \quad (12)$$

25 où  $cxl[n,m]$  désigne le coefficient cepstral lissé d'ordre  $m$  pour la trame  $n$ .

A l'étape 75, l'index d'harmonique  $k$  est initialisé à 1. Pour initialiser le calcul de la phase minimale affectée à l'harmonique  $k$ , la phase  $\varphi(k)$  et l'index cepstral  $m$  sont initialisés respectivement à 0 et 1 à l'étape 76. A l'étape 77, le module 57 ajoute à la phase  $\varphi(k)$  la quantité  $-2 \cdot cxl[n,m] \cdot \sin(2\pi m k F_0 / F_e)$ .

30 L'index cepstral  $m$  est incrémenté à l'étape 78 et comparé à NCS à l'étape 79.

Les étapes 77 et 78 sont répétées tant que  $m < \text{NCS}$ . Quand  $m = \text{NCS}$ , le calcul de la phase minimale est terminé pour l'harmonique  $k$ , et l'index  $k$  est incrémenté à l'étape 80. Le calcul de phases minimales 76-79 est renouvelé pour l'harmonique suivante tant que  $k.F_0 < F_e/2$  (test 81).

5 Dans l'exemple de réalisation selon la figure 8, le module 54 tient compte d'une phase constante sur la largeur de chaque raie spectrale, égale à la phase minimale  $\phi(k)$  fournie pour l'harmonique correspondante  $k$  par le module 57.

L'estimation  $X_v$  de la composante harmonique est synthétisée par  
10 sommation de raies spectrales positionnées aux fréquences harmoniques de la fréquence fondamentale  $F_0$ . Lors de cette synthèse, on peut positionner les raies spectrales sur l'axe des fréquences avec une résolution supérieure à la résolution de la transformée de Fourier. Pour cela, on précalcule une fois pour toutes une raie spectrale de référence selon la résolution supérieure. Ce calcul  
15 peut consister en une transformée de Fourier de la fenêtre d'analyse  $f_A$  avec une taille de transformée de 16384 points, procurant une résolution de 0,5 Hz par point. La synthèse de chaque raie harmonique est alors effectuée par le module 54 en positionnant sur l'axe des fréquences la raie de référence à haute résolution, et en sous-échantillonnant cette raie spectrale de référence  
20 pour se ramener à la résolution de 16,625 Hz de la transformée de Fourier sur 512 points. Ceci permet de positionner avec précision la raie spectrale.

Pour la détermination de l'enveloppe inférieure, le module de TFR 85 du décodeur de la figure 8 reçoit les NCI coefficients cepstraux quantifiés  $cx\_inf\_q$  d'ordres 0 à  $\text{NCI} - 1$ , et il les complète avantageusement par les  
25  $\text{NCS} - \text{NCI}$  coefficients cepstraux  $cx\_sup\_q$  d'ordre  $\text{NCI}$  à  $\text{NCS} - 1$  représentant l'enveloppe supérieure. En effet, on peut estimer en première approximation que les variations rapides de l'enveloppe inférieure comprimée sont bien reproduites par celles de l'enveloppe supérieure comprimée. Dans une autre réalisation, le module de TFR 85 pourrait ne considérer que les NCI  
30 paramètres cepstraux  $cx\_inf\_q$ .

Le module 86 convertit l'échelle de fréquences de manière réciproque de la conversion opérée par le module 32 du codeur, afin de restituer l'estimation  $LX\_inf$  de l'enveloppe inférieure comprimée, soumise au module de décompression spectrale 87. En sortie du module 87, le décodeur dispose



d'une enveloppe inférieure  $X_{\text{inf}}$  comportant les valeurs du module du spectre dans les vallées situées entre les pics harmoniques.

Cette enveloppe  $X_{\text{inf}}$  va moduler le spectre d'une trame de bruit dont la phase est traitée en fonction du vecteur de pondération quantifié  $E_{\text{mix}}$  extrait par le module 49. Un générateur 88 délivre une trame de bruit normalisé dont les segments de 4 ms sont pondérés dans un module 89 conformément aux composantes normalisées du vecteur  $E_{\text{mix}}$  fourni par le module 49 pour la trame courante. Ce bruit est un bruit blanc filtré passe-haut pour tenir compte du faible niveau qu'a en principe la composante non-voisée aux basses fréquences. A partir du bruit pondéré en énergie, le module 90 forme des trames de  $2N_s = 512$  échantillons en appliquant la fenêtre d'analyse  $f_A$ , l'insertion de 256 échantillons à zéro et la permutation circulaire pour la compensation de phase conformément à ce qui a été expliqué en référence à la figure 3. La transformée de Fourier de la trame résultante est calculée par le module TFR 91.

L'estimation spectrale  $X_{uv}$  de la composante non-harmonique est déterminée par le module de synthèse spectrale 92 qui effectue une pondération fréquence par fréquence. Cette pondération consiste à multiplier chaque valeur spectrale complexe fournie par le module de TFR 91 par la valeur de l'enveloppe inférieure  $X_{\text{inf}}$  obtenue pour la même fréquence par le module de décompression spectrale 87.

Les estimations spectrales  $X_v$ ,  $X_{uv}$  des composantes harmonique (voisée dans le cas d'un signal de parole) et non-harmonique (ou non-voisée) sont combinées par un module de mixage 95 contrôlé par un module 96 d'analyse du degré d'harmonicité (ou de voisement) du signal.

L'organisation de ces modules 95, 96 est illustrée par la figure 10. Le module d'analyse 96 comporte une unité 97 d'estimation d'un degré de voisement  $W$  dépendant de la fréquence, à partir duquel sont calculés quatre gains dépendant de la fréquence, à savoir deux gains  $g_v$ ,  $g_{uv}$  contrôlant l'importance relative des composantes harmonique et non-harmonique dans le signal synthétisé, et deux gains  $g_{v_\varphi}$ ,  $g_{uv_\varphi}$  utilisés pour bruiteur la phase de la composante harmonique.

Le degré de voisement  $W(i)$  est une valeur à variation continue comprise entre 0 et 1 déterminée pour chaque index de fréquence  $i$  ( $0 \leq i < N$ ) en fonction de l'enveloppe supérieure  $X_{\text{sup}}(i)$  et de l'enveloppe inférieure

$X_{\text{inf}}(i)$  obtenues pour cette fréquence  $i$  par les modules de décompression 53, 87. Le degré de voisement  $W(i)$  est estimé par l'unité 97 pour chaque index de fréquence  $i$  correspondant à une harmonique de la fréquence fondamentale  $F_0$ ,

à savoir  $i = \left\lfloor 2Nk \frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$  pour  $k = 1, 2, \dots$ , par une fonction croissante du

- 5 rapport entre l'enveloppe supérieure  $X_{\text{sup}}$  et l'enveloppe inférieure  $X_{\text{inf}}$  à cette fréquence, par exemple selon la formule :

$$W(i) = \min \left\{ 1, \frac{10 \cdot \log_{10} [X_{\text{sup}}(i)/X_{\text{inf}}(i)]}{V_{\text{th}}(F_0)} \right\} \quad (13)$$

- Le seuil  $V_{\text{th}}(F_0)$  correspond à la dynamique moyenne calculée sur un spectre synthétique purement voisé à la fréquence fondamentale. Il est  
10 avantageusement choisi dépendant de la fréquence fondamentale  $F_0$ .

Le degré de voisement  $W(i)$  pour une fréquence autre que les fréquences harmoniques est obtenu simplement comme étant égal à celui estimé pour l'harmonique la plus proche.

- Le gain  $g_v(i)$ , qui dépend de la fréquence, est obtenu en appliquant une  
15 fonction non-linéaire au degré de voisement  $W(i)$  (bloc 98). Cette fonction non-linéaire a par exemple la forme représentée sur la figure 11 :

$$\begin{aligned} g_v(i) &= 0 \quad \text{si } 0 \leq W(i) \leq W1 \\ g_v(i) &= \frac{W(i) - W1}{W2 - W1} \quad \text{si } W1 < W(i) < W2 \\ g_v(i) &= 1 \quad \text{si } W2 \leq W(i) \leq 1 \end{aligned} \quad (14)$$

- 20 les seuils  $W1$ ,  $W2$  étant tels que  $0 < W1 < W2 < 1$ . Le gain  $g_{uv}$  peut être calculé de manière semblable au gain  $g_v$  (la somme des deux gains  $g_v$ ,  $g_{uv}$  étant constante, par exemple égale à 1), ou déduit simplement de celui-ci par la relation  $g_{uv}(i) = 1 - g_v(i)$ , comme schématisé par le soustracteur 99 sur la figure 10.

- 25 Il est intéressant de pouvoir bruite la phase de la composante harmonique du signal à une fréquence donnée si l'analyse du degré de voisement montre que le signal est plutôt de type non-harmonique à cette fréquence. Pour cela, la phase  $\phi_v$  de la composante harmonique mixée est le résultat d'une combinaison linéaire des phases  $\phi_v$ ,  $\phi_{uv}$  des composantes  
30 harmonique et non-harmonique  $X_v$ ,  $X_{uv}$  synthétisées par les modules 54, 92.

Les gains  $g_{v\_φ}$ ,  $g_{uv\_φ}$  respectivement appliqués à ces phases sont calculés à partir du degré de voisement  $W$  et pondérés également en fonction de l'index de fréquence  $i$ , étant donné que le bruitage de la phase n'est véritablement utile qu'au-delà d'une certaine fréquence.

- 5 Un premier gain  $g_{v1\_φ}$  est calculé en appliquant une fonction non-linéaire au degré de voisement  $W(i)$ , comme schématisé par le bloc 100 sur la figure 10. Cette fonction non-linéaire peut avoir la forme représentée sur la figure 12 :

$$\begin{aligned} g_{v1\_φ}(i) &= G1 \quad \text{si } 0 \leq W(i) \leq W3 \\ 10 \quad g_{v1\_φ}(i) &= G1 + (1 - G1) \frac{W(i) - W3}{W4 - W3} \quad \text{si } W3 < W(i) < W4 \\ g_{v1\_φ}(i) &= 1 \quad \text{si } W4 \leq W(i) \leq 1 \end{aligned} \quad (15)$$

les seuils  $W3$  et  $W4$  étant tels que  $0 < W3 < W4 < 1$ , et le gain minimal  $G1$  étant compris entre 0 et 1.

- 15 Un multiplieur 101 multiplie pour chaque fréquence d'index  $i$  le gain  $g_{v1\_φ}$  par un autre gain  $g_{v2\_φ}$  dépendant seulement de l'index de fréquence  $i$ , pour former le gain  $g_{v\_φ}(i)$ . Le gain  $g_{v2\_φ}(i)$  dépend non-linéairement de l'index de fréquence  $i$ , par exemple comme indiqué sur la figure 13 :

$$\begin{aligned} g_{v2\_φ}(i) &= 1 \quad \text{si } 0 \leq i \leq i1 \\ 20 \quad g_{v2\_φ}(i) &= 1 - (1 - G2) \frac{i - i1}{i2 - i1} \quad \text{si } i1 < i < i2 \\ g_{v2\_φ}(i) &= G2 \quad \text{si } i2 \leq i \leq 1 \end{aligned} \quad (16)$$

les index  $i1$  et  $i2$  étant tels que  $0 < i1 < i2 \leq N$ , et le gain minimal  $G2$  étant compris entre 0 et 1. Le gain  $g_{uv\_φ}(i)$  peut être calculé simplement comme étant égal à  $1 - g_{v\_φ}(i) = 1 - g_{v1\_φ}(i) \cdot g_{v2\_φ}(i)$  (soustracteur 102 de la figure 10).

- 25 Le spectre complexe  $Y$  du signal synthétisé est produit par le module de mixage 95, qui réalise la relation de mixage suivante, pour  $0 \leq i < N$  :

$$Y(i) = g_v(i) \cdot |X_v(i)| \cdot \exp[j \phi_v(i)] + g_{uv}(i) \cdot X_{uv}(i) \quad (17)$$

avec  $\phi_v(i) = g_{v\_φ}(i) \cdot \phi_v(i) + g_{uv\_φ}(i) \cdot \phi_{uv}(i) \quad (18)$

- où  $\phi_v(i)$  désigne l'argument du nombre complexe  $X_v(i)$  fourni par le module 54  
30 pour la fréquence d'index  $i$  (bloc 104 de la figure 10), et  $\phi_{uv}(i)$  désigne

l'argument du nombre complexe  $X_{uv}(i)$  fourni par le module 92 (bloc 105 de la figure 10). Cette combinaison est réalisée par les multiplieurs 106-110 et les additionneurs 111-112 représentés sur la figure 10.

Le spectre mixé  $Y(i)$  pour  $0 \leq i < 2N$  (avec  $Y(2N-1-i) = Y(i)$ ) est ensuite transformé dans le domaine temporel par le module de TFRI 115 (figure 8). On ne retient que les  $N/2 = 128$  premiers et les  $N/2 = 128$  derniers échantillons de la trame de  $2N = 512$  échantillons produite par le module 115, et on applique la permutation circulaire inverse de celle illustrée par la figure 3 pour obtenir la trame synthétisée de  $N = 256$  échantillons pondérés par la fenêtre d'analyse  $f_A$ .

Les trames successivement obtenues de cette manière sont finalement traitées par le module de synthèse temporelle 116 qui forme le signal audio décodé  $\hat{x}$ .

Le module de synthèse temporelle 116 effectue une somme à recouvrement de trames modifiées par rapport à celles successivement évaluées en sortie du module 115. La modification peut être vue en deux étapes illustrées respectivement par les figures 14 et 15.

La première étape (figure 14) consiste à multiplier chaque trame 2' délivrée par le module de TFRI 115 par une fenêtre  $1/f_A$  inverse de la fenêtre d'analyse  $f_A$  employée par le module 1 du codeur. Les échantillons de la trame 2" qui en résultent sont donc pondérés uniformément.

La seconde étape (figure 15) consiste à multiplier les échantillons de cette trame 2" par une fenêtre de synthèse  $f_S$  vérifiant les propriétés suivantes :

$$f_S(N-L+i) + f_S(i) = A \quad \text{pour } 0 \leq i < L \quad (19)$$

$$f_S(i) = A \quad \text{pour } L \leq i < N-L \quad (20)$$

où  $A$  désigne une constante positive arbitraire, par exemple  $A = 1$ . La fenêtre de synthèse  $f_S(i)$  croît progressivement de 0 à  $A$  pour  $i$  allant de 0 à  $L$ . C'est par exemple une demi-sinusoïde surélevée :

$$f_S(i) = \frac{A}{2} \cdot (1 - \cos[(i+1/2)\pi/L]) \quad \text{pour } 0 \leq i < L \quad (21)$$

Après avoir repondéré chaque trame 2" par la fenêtre de synthèse  $f_S$ , le module 116 positionne les trames successives avec leurs décalages temporels de  $M = 160$  échantillons et leurs recouvrements temporels de  $L = 96$  échantillons, puis il effectue la somme des trames ainsi positionnées dans le temps. Du fait des propriétés (19) et (20) de la fenêtre de synthèse  $f_S$ , chaque

échantillon du signal audio décodé  $\hat{x}$  ainsi obtenu est affecté d'un poids global uniforme, égal à A. Ce poids global provient de la contribution d'une trame unique si l'échantillon a dans cette trame un rang  $i$  tel que  $L \leq i < N - L$ , et comporte les contributions sommées de deux trames successives si  $0 \leq i < L$  ou  $N - L \leq i < N$ .

On peut ainsi effectuer la synthèse temporelle de façon simple même si, comme dans le cas considéré, le recouvrement  $L$  entre deux trames successives est plus petit que la moitié de la taille  $N$  de ces trames.

Les deux étapes exposées ci-dessus pour la modification des trames de signal peuvent être fusionnées en une seule étape. Il suffit de précalculer une fenêtre composée  $f_C(i) = f_S(i)/f_A(i)$ , et de multiplier simplement les trames  $2'$  de  $N = 256$  échantillons délivrées par le module 115 par la fenêtre composée  $f_C$  avant d'effectuer la sommation à recouvrement.

La figure 16 montre l'allure de la fenêtre composée  $f_C$  dans le cas où la fenêtre d'analyse  $f_A$  est une fenêtre de Hamming et la fenêtre de synthèse  $f_S$  a la forme donnée par les relations (19) à (21).

D'autres formes de la fenêtre de synthèse  $f_S$  vérifiant les relations (19) et (20) peuvent être employées. Dans la variante de la figure 17, c'est une fonction affine par morceaux définie par :

$$f_S(i) = A.i/L \quad \text{pour } 0 \leq i < L \quad (22)$$

Afin d'améliorer la qualité de codage du signal audio, le codeur de la figure 1 peut augmenter la cadence de formation et d'analyse des trames, afin de transmettre davantage de paramètres de quantification au décodeur. Dans la structure de trame représentée sur la figure 2, une trame de  $N = 256$  échantillons (32 ms) est formée toutes les 20 ms. Ces trames de 256 échantillons pourraient être formées à une cadence supérieure, par exemple de 10 ms, deux trames successives ayant alors un décalage de  $M/2 = 80$  échantillons et un recouvrement de 176 échantillons.

Dans ces conditions, on peut transmettre les jeux complets de paramètres de quantification  $iF$ ,  $icxs$ ,  $icxi$ ,  $iEm$  pour seulement un sous-ensemble des trames, et transmettre pour les autres trames des paramètres permettant d'effectuer une interpolation adéquate au niveau du décodeur. Dans l'exemple envisagé ci-dessus, le sous-ensemble pour lequel des jeux de paramètres complets sont transmis peut être constitué par les trames de rang

entier  $n$ , dont la périodicité est de  $M/F_e = 20$  ms, et les trames pour lesquelles une interpolation est effectuée peuvent être celles de rang demi-entier  $n + 1/2$  qui sont décalées de 10 ms par rapport aux trames du sous-ensemble.

Dans la réalisation illustrée par la figure 18, les notations  $cx\_q[n-1]$  et  $cx\_q[n]$  désignent des vecteurs cepstraux quantifiés déterminés, pour deux trames successives de rang entier, par le module de quantification 18 et/ou par le module de quantification 34. Ces vecteurs comprennent par exemple quatre coefficients cepstraux consécutifs chacun. Ils pourraient également comprendre davantage de coefficients cepstraux.

Un module 120 effectue une interpolation de ces deux vecteurs cepstraux  $cx\_q[n-1]$  et  $cx\_q[n]$ , afin d'estimer une valeur intermédiaire  $cx\_i[n-1/2]$ . L'interpolation effectuée par le module 120 peut être une simple moyenne arithmétique des vecteurs  $cx\_q[n-1]$  et  $cx\_q[n]$ . En variante, le module 120 pourrait appliquer une formule d'interpolation plus sophistiquée, par exemple polynomiale, en se fondant également sur les vecteurs cepstraux obtenus pour des trames antérieures à la trame  $n-1$ . D'autre part, si plus d'une trame interpolée est intercalée entre deux trames consécutives de rang entier, l'interpolation tient compte de la position relative de chaque trame interpolée.

A l'aide des moyens précédemment décrits, le codeur calcule également les coefficients cepstraux  $cx[n-1/2]$  relatifs à la trame de rang demi-entier. Dans le cas de l'enveloppe supérieure, ces coefficients cepstraux sont ceux fournis par le module de TFRI 13 après post-liffrage 15 (par exemple avec les mêmes coefficients de post-liffrage que pour la trame précédente  $n-1$ ) et normalisation 16. Dans le cas de l'enveloppe inférieure, les coefficients cepstraux  $cx[n-1/2]$  sont ceux délivrés par le module de TFRI 33.

Un soustracteur 121 forme la différence  $ecx[n-1/2]$  entre les coefficients cepstraux  $cx[n-1/2]$  calculés pour la trame de rang demi-entier et les coefficients  $cx\_i[n-1/2]$  estimés par interpolation. Cette différence est fournie à un module de quantification 122 qui adresse des index de quantification  $icx[n-1/2]$  au multiplexeur de sortie 6 du codeur. Le module 122 fonctionne par exemple par quantification vectorielle des erreurs d'interpolation  $ecx[n-1/2]$  successivement déterminées pour les trames de rang demi-entier.

Cette quantification de l'erreur d'interpolation peut être effectuée par le codeur pour chacun des NCS + NCI coefficients cepstraux utilisés par le décodeur, ou seulement pour certains d'entre eux, typiquement ceux d'ordres

les plus petits.

Les moyens correspondants du décodeur sont illustrés par la figure 19. Le décodeur fonctionne essentiellement comme celui décrit en référence à la figure 8 pour déterminer les trames de signal de rang entier. Un module d'interpolation 124 identique au module 120 du codeur estime les coefficients intermédiaires  $cx\_i[n-1/2]$  à partir des coefficients quantifiés  $cx\_q[n-1]$  et  $cx\_q[n]$  fournis par le module 47 et/ou le module 48 à partir des index  $icxs$ ,  $icxi$  extraits du flux  $\Phi$ . Un module d'extraction de paramètres 125 reçoit l'index de quantification  $icx[n-1/2]$  depuis le démultiplexeur d'entrée 45 du décodeur, et en déduit l'erreur d'interpolation quantifiée  $ecx\_q[n-1/2]$  à partir du même dictionnaire de quantification que celui utilisé par le module 122 du codeur. Un additionneur 126 fait la somme des vecteurs cepstraux  $cx\_i[n-1/2]$  et  $ecx\_q[n-1/2]$  afin de fournir les coefficients cepstraux  $cx[n-1/2]$  qui seront utilisés par le décodeur (modules 51-57, 95, 96, 115 et/ou modules 85-87, 92, 95, 96, 115) pour former la trame interpolée de rang  $n-1/2$ .

Si certains seulement des coefficients cepstraux ont fait l'objet d'une quantification d'erreur d'interpolation, les autres sont déterminés par le décodeur par une interpolation simple, sans correction.

Le décodeur peut également interpoler les autres paramètres  $F_0$ ,  $Emix$  utilisés pour synthétiser les trames de signal. La fréquence fondamentale  $F_0$  peut être interpolée linéairement, soit dans le domaine temporel, soit (de préférence) directement dans le domaine fréquentiel. Pour l'interpolation éventuelle du vecteur de pondération énergétique  $Emix$ , il convient d'effectuer l'interpolation après dénormalisation et en tenant compte bien entendu des décalages temporels entre trames.

Il est à noter qu'il est particulièrement avantageux, pour interpoler la représentation des enveloppes spectrales, d'effectuer cette interpolation dans le domaine cepstral. Contrairement à une interpolation effectuée sur d'autres paramètres, tels que les coefficients LSP (« Line Spectrum Pairs »), l'interpolation linéaire des coefficients cepstraux correspond à l'interpolation linéaire des amplitudes spectrales comprimées.

Dans la variante représentée sur la figure 20, le codeur utilise les vecteurs cepstraux  $cx\_q[n]$ ,  $cx\_q[n-1]$ , ...,  $cx\_q[n-r]$  et  $cx\_q[n-1/2]$  calculés pour les dernières trames passées ( $r \geq 1$ ) pour identifier un filtre interpolateur optimal qui, lorsqu'on lui soumet les vecteurs cepstraux quantifiés  $cx\_q[n-r]$ ,

...,  $cx\_q[n]$  relatifs aux trames de rang entier, délivre un vecteur cepstral interpolé  $cx\_i[n-1/2]$  qui présente une distance minimale avec le vecteur  $cx[n-1/2]$  calculé pour la dernière trame de rang demi-entier.

Dans l'exemple représenté sur la figure 20, ce filtre interpolateur 128 est présent dans le codeur, et un soustracteur 129 retranche sa sortie  $cx\_i[n-1/2]$  du vecteur cepstral calculé  $cx[n-1/2]$ . Un module de minimisation 130 détermine le jeu de paramètres  $\{P\}$  du filtre interpolateur 128, pour lequel l'erreur d'interpolation  $ecx[n-1/2]$  délivrée par le soustracteur 129 présente une norme minimale. Ce jeu de paramètres  $\{P\}$  est adressé à un module de quantification 131 qui fournit un index de quantification correspondant  $iP$  au multiplexeur de sortie 6 du codeur.

En fonction du débit alloué dans le flux  $\Phi$  aux index de quantification des paramètres  $\{P\}$  définissant le filtre interpolateur optimal 128, on pourra adopter une quantification plus ou moins fine de ces paramètres, ou une forme plus ou moins élaborée du filtre interpolateur, ou encore prévoir plusieurs filtres interpolateurs quantifiés de manière distincte pour différents vecteurs de coefficients cepstraux.

Dans une réalisation simple, le filtre interpolateur 128 est linéaire, avec  $r = 1$  :

$$cx\_i[n-1/2] = \rho \cdot cx\_q[n-1] + (1-\rho) \cdot cx\_q[n] \quad (23)$$

et le jeu de paramètres  $\{P\}$  se limite au coefficient  $\rho$  compris entre 0 et 1.

A partir des index  $iP$  de quantification des paramètres  $\{P\}$  obtenus dans le flux binaire  $\phi$ , le décodeur reconstruit le filtre interpolateur 128 (aux erreurs de quantification près), et traite les vecteurs spectraux  $cx\_q[n-r]$ , ...,  $cx\_q[n]$  afin d'estimer les coefficients cepstraux  $cx[n-1/2]$  utilisés pour synthétiser les trames de rang demi-entier.

De façon générale, le décodeur peut utiliser une méthode d'interpolation simple (sans transmission de paramètres de la part du codeur pour les trames de rang demi-entier), une méthode d'interpolation avec prise en compte d'une erreur d'interpolation quantifiée (selon les figures 17 et 18), ou une méthode d'interpolation avec un filtre interpolateur optimal (selon la figure 19) pour évaluer les trames de rang demi-entier en plus des trames de rang entier évaluées directement comme expliqué en référence aux figures 8 à 13. Le module 116 de synthèse temporelle peut alors combiner l'ensemble de ces trames évaluées pour former le signal synthétisé  $\hat{x}$  de la manière expliquée ci-



après en référence aux figures 14, 21 et 22.

Comme dans la méthode de synthèse temporelle précédemment décrite, le module 116 effectue une somme à recouvrement de trames modifiées par rapport à celles successivement évaluées en sortie du module 115, et cette modification peut être vue en deux étapes dont la première est identique à celle précédemment décrite en référence à la figure 14 (diviser les échantillons de la trame 2' par la fenêtre d'analyse  $f_A$ ).

La seconde étape (figure 21) consiste à multiplier les échantillons de la trame renormalisée 2'' par une fenêtre de synthèse  $f'_S$  vérifiant les propriétés suivantes :

$$f'_S(i) = 0 \quad \text{pour } 0 \leq i < N/2 - M/p \quad \text{et} \quad N/2 + M/p \leq i < N \quad (24)$$

$$f'_S(i) + f'_S(i+M/p) = A \quad \text{pour } N/2 - M/p \leq i < N/2 \quad (25)$$

où  $A$  désigne une constante positive arbitraire, par exemple  $A = 1$ , et  $p$  est l'entier tel que le décalage temporel entre les trames successives (calculées directement et interpolées) soit de  $M/p$  échantillons, soit  $p = 2$  dans l'exemple décrit. La fenêtre de synthèse  $f'_S(i)$  croît progressivement pour  $i$  allant de  $N/2 - M/p$  à  $N/2$ . C'est par exemple une sinusoïde surélevée sur l'intervalle  $N/2 - M/p \leq i < N/2 + M/p$ . En particulier, la fenêtre de synthèse  $f'_S$  peut être, sur cet intervalle, une fenêtre de Hamming (comme représenté sur la figure 21) ou une fenêtre de Hanning.

La figure 21 montre les trames successives 2'' repositionnées dans le temps par le module 116. Les hachures indiquent les portions éliminées des trames (fenêtre de synthèse à 0). On voit qu'en effectuant la somme à recouvrement des échantillons des trames successives, la propriété (25) assure une pondération homogène des échantillons du signal synthétisé.

Comme dans la méthode de synthèse illustrée par les figures 14 et 15, la procédure de pondération des trames obtenues par transformée de Fourier inverse des spectres  $Y$  peut être effectuée en une seule étape, avec une fenêtre composée  $f'_C(i) = f'_S(i)/f_A(i)$ . La figure 22 montre la forme de la fenêtre composée  $f'_C$  dans le cas où les fenêtres  $f_A$  et  $f'_S$  sont de type Hamming.

Comme la méthode de synthèse temporelle illustrée par les figures 14 à 17, celle illustrée par les figures 14, 21 et 22 permet de prendre en compte

un recouvrement  $L$  entre deux trames d'analyse (pour lesquelles l'analyse est effectuée de façon complète) plus petit que la moitié que la taille  $N$  de ces trames. De façon générale, cette dernière méthode est applicable lorsque les trames d'analyse successives présentent des décalages temporels mutuels  $M$  de plus de  $N/2$  échantillons (même éventuellement de plus de  $N$  échantillons si un très bas débit est requis), l'interpolation conduisant à un ensemble de trames dont les décalages temporels mutuels sont de moins de  $N/2$  échantillons.

Les trames interpolées peuvent faire l'objet d'une transmission réduite de paramètres de codage, comme décrit précédemment, mais cela n'est pas obligatoire. Ce mode de réalisation permet de conserver un intervalle  $M$  relativement grand entre deux trames d'analyse, et donc de limiter le débit de transmission requis, tout en limitant les discontinuités susceptibles d'apparaître en raison de la taille de cet intervalle par rapport aux échelles de temps typiques des variations des paramètres du signal audio, notamment les coefficients cepstraux et la fréquence fondamentale.

## REVENDICATIONS

1. Procédé d'analyse d'un signal audio (x) traité par trames successives de N échantillons, dans lequel on pondère les échantillons de chaque trame par une fenêtre d'analyse ( $f_A$ ) de type Hamming, Hanning,  
5 Kaiser ou analogue, on calcule un spectre du signal audio en transformant chaque trame d'échantillons pondérés dans le domaine fréquentiel, et on traite le spectre du signal audio pour délivrer des paramètres ( $cx\_sup$ ,  $cx\_inf$ ,  $Emix$ ) de synthèse d'un signal dérivé du signal audio analysé, caractérisé en ce que les trames successives comportent une alternance de trames pour lesquelles  
10 on délivre des jeux complets de paramètres de synthèse et de trames pour lesquelles on délivre des jeux incomplets de paramètres de synthèse, et en ce que les trames successives pour lesquelles on délivre des jeux complets de paramètres de synthèse présentent des recouvrements mutuels de moins de  $N/2$  échantillons.
- 15 2. Procédé selon la revendication 1, dans lequel les jeux incomplets de paramètres de synthèse incluent des données ( $icx[n-1/2]$ ) représentant une erreur ( $ecx[n-1/2]$ ) d'interpolation d'au moins un des paramètres de synthèse.
3. Procédé selon la revendication 1, dans lequel les jeux incomplets de paramètres de synthèse incluent des données ( $iP$ ) représentant un filtre (128)  
20 d'interpolation d'au moins un des paramètres de synthèse.
4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel le traitement du spectre du signal audio (x) comporte une extraction de paramètres de codage ( $cx\_sup$ ,  $cx\_inf$ ,  $Emix$ ) en vue de la transmission et/ou du stockage du signal audio codé.
- 25 5. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel le traitement du spectre du signal audio (x) comporte un débruitage par soustraction spectrale.
6. Dispositif de traitement audio, comprenant des moyens d'analyse pour exécuter un procédé selon l'une quelconque des revendications 1 à 5.

7. Procédé de synthèse d'un signal audio, dans lequel on obtient des estimations spectrales successives ( $Y$ ) correspondant respectivement à des trames de  $N$  échantillons du signal audio pondérés par une fenêtre d'analyse ( $f_A$ ), les trames successives présentant des recouvrements mutuels de  $L$  échantillons, on évalue chaque trame du signal audio en transformant les estimations spectrales dans le domaine temporel, et on combine les trames évaluées pour former le signal synthétisé ( $\hat{x}$ ), caractérisé en ce qu'on modifie chaque trame évaluée en lui appliquant un traitement correspondant à une division par ladite fenêtre d'analyse ( $f_A$ ) et à une multiplication par une fenêtre de synthèse ( $f_S$ ), et on forme le signal synthétisé comme une somme à recouvrement des trames modifiées, et en ce que, le nombre  $L$  étant plus petit que  $N/2$  et les échantillons d'une trame ayant des rangs  $i$  numérotés de 0 à  $N-1$ , la fenêtre de synthèse  $f_S(i)$  vérifie  $f_S(N-L+i) + f_S(i) = A$  pour  $0 \leq i < L$ , et est égale à  $A$  pour  $L \leq i < N-L$ ,  $A$  étant une constante positive.
8. Procédé selon la revendication 7, dans lequel la fenêtre de synthèse  $f_S(i)$  croît de 0 à  $A$  pour  $i$  allant de 0 à  $L$ .
9. Procédé selon la revendication 8, dans lequel la fenêtre de synthèse  $f_S(i)$  pour  $0 \leq i < L$  est une demi-sinusoïde surélevée.
10. Procédé de synthèse d'un signal audio, dans lequel on évalue un ensemble de trames recouvrantes successives de  $N$  échantillons du signal audio pondérés par une fenêtre d'analyse ( $f_A$ ), en transformant dans le domaine temporel des estimations spectrales ( $Y$ ) correspondant respectivement auxdites trames, et on combine les trames évaluées pour former le signal synthétisé ( $\hat{x}$ ), caractérisé en ce que, pour un sous-ensemble des trames évaluées, les estimations spectrales sont obtenues en traitant des paramètres de synthèse ( $cx\_sup\_q$ ,  $cx\_inf\_q$ ,  $E_{mix}$ ) respectivement associés aux trames dudit sous-ensemble tandis que, pour les trames ne faisant pas partie du sous-ensemble, les estimations spectrales sont obtenues avec une interpolation d'une partie au moins des paramètres de synthèse, en ce que les trames successives dudit sous-ensemble présentent des décalages temporels mutuels de  $M$  échantillons, le nombre  $M$  étant plus grand que  $N/2$ , tandis que les trames successives dudit ensemble présentent des décalages temporels

mutuels de  $M/p$  échantillons,  $p$  étant un entier plus grand que 1, en ce qu'on modifie chaque trame évaluée en lui appliquant un traitement correspondant à une division par ladite fenêtre d'analyse ( $f_A$ ) et à une multiplication par une fenêtre de synthèse ( $f'_S$ ), et on forme le signal synthétisé comme une somme à recouvrement des trames modifiées, et en ce que, les échantillons d'une trame ayant des rangs  $i$  numérotés de 0 à  $N-1$ , la fenêtre de synthèse  $f'_S(i)$  a un support limité aux rangs  $i$  allant de  $N/2 - M/p$  à  $N/2 + M/p$  et vérifie  $f'_S(i) + f'_S(i+M/p) = A$  pour  $N/2 - M/p \leq i < N/2$ ,  $A$  étant une constante positive.

11. Procédé selon la revendication 10, dans lequel la fenêtre de synthèse  $f'_S(i)$  croît pour  $i$  allant de  $N/2 - M/p$  à  $N/2$ .

12. Procédé selon la revendication 11, dans lequel la fenêtre de synthèse  $f'_S(i)$  pour  $N/2 - M/p \leq i < N/2 + M/p$  est une sinusoïde surélevée.

13. Procédé selon l'une quelconque des revendications 10 à 12, dans lequel des données ( $icx\_q[n-1/2]$ ) représentant une erreur d'interpolation ( $ecx\_q[n-1/2]$ ) sont associées aux trames ne faisant pas partie dudit sous-ensemble, et sont utilisées pour corriger au moins un des paramètres de synthèse interpolés ( $cx\_i[n-1/2]$ ).

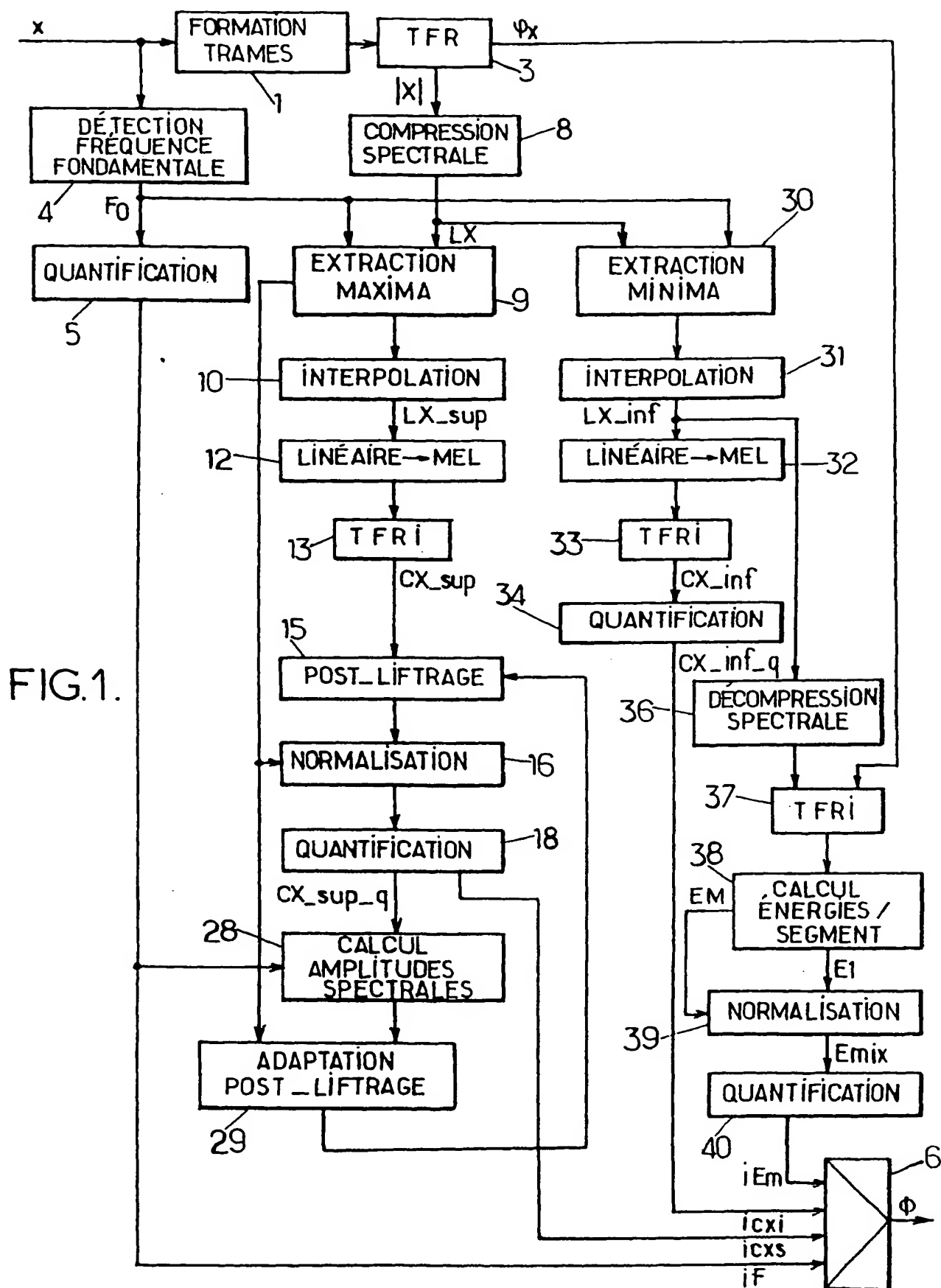
14. Procédé selon l'une quelconque des revendications 10 à 12, dans lequel des données ( $iP$ ) représentant un filtre interpolateur (128) sont associées aux trames ne faisant pas partie dudit sous-ensemble, et sont utilisées pour interpoler au moins un des paramètres de synthèse.

15. Procédé selon l'une quelconque des revendications 10 à 14, dans lequel les paramètres de synthèse comprennent des coefficients cepstraux ( $cx[n]$ ) soumis à l'interpolation.

16. Dispositif de traitement audio, comprenant des moyens de synthèse pour exécuter un procédé selon l'une quelconque des revendications 7 à 15.

**THIS PAGE BLANK (USPT**

1/11



2 PAGE BLANK (USPTO)



2/11

FIG. 2

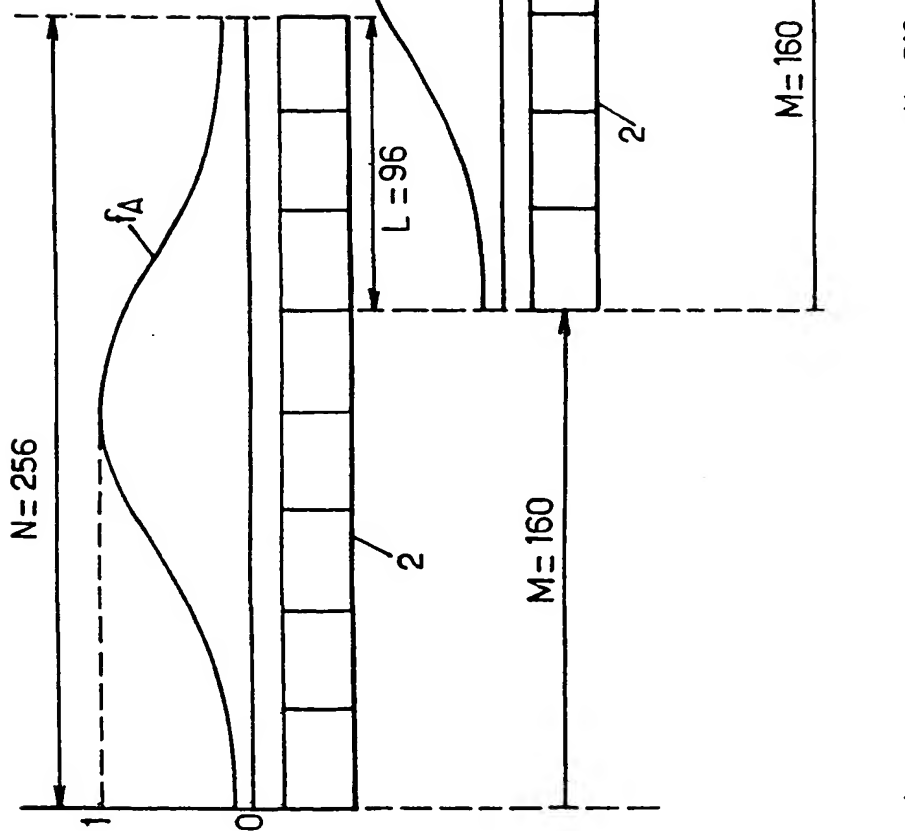
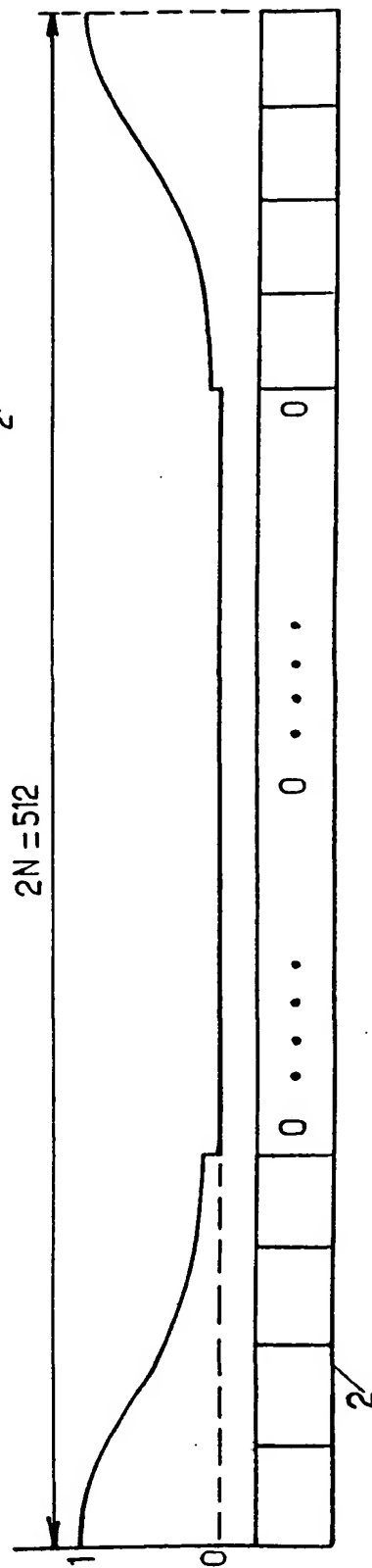


FIG. 3.



**THIS PAGE BLANK (USPTO)**

3/11

FIG.4.

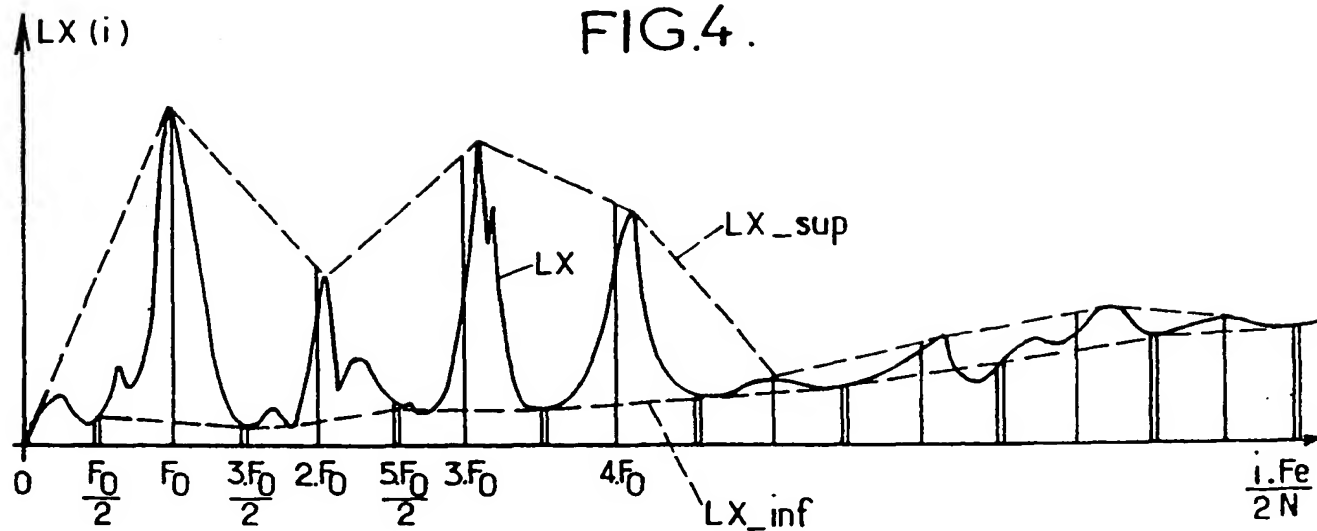
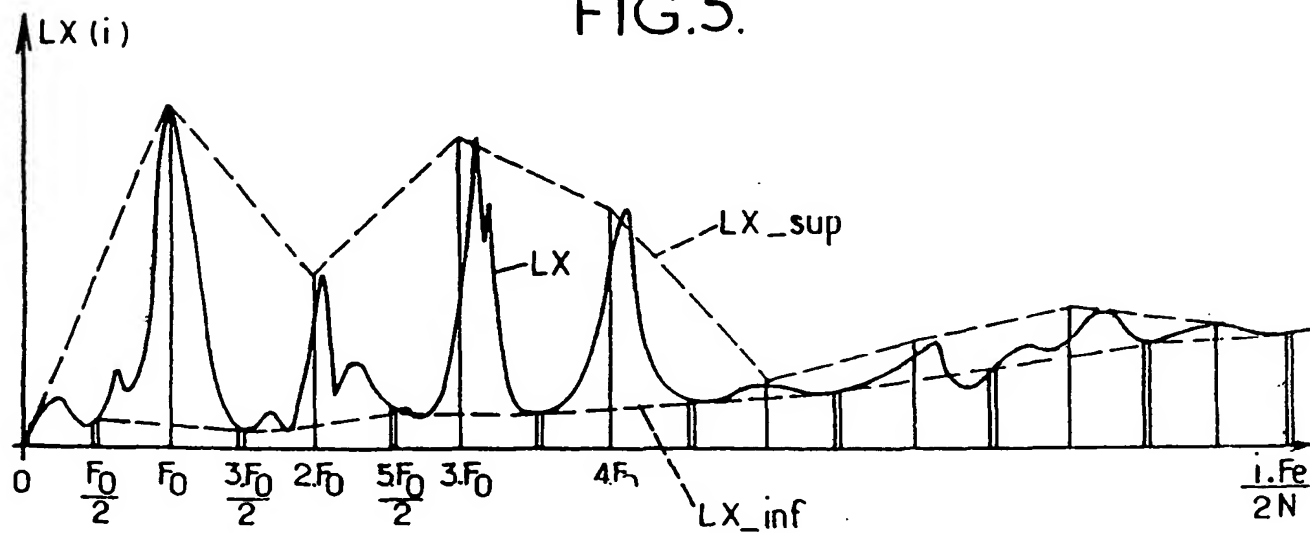


FIG.5.



**THIS PAGE BLANK (USPTO)**

4/11

FIG. 6.

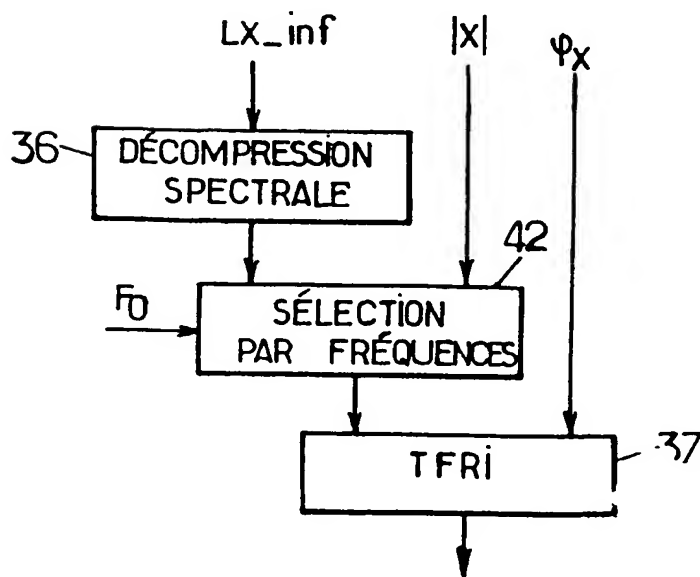
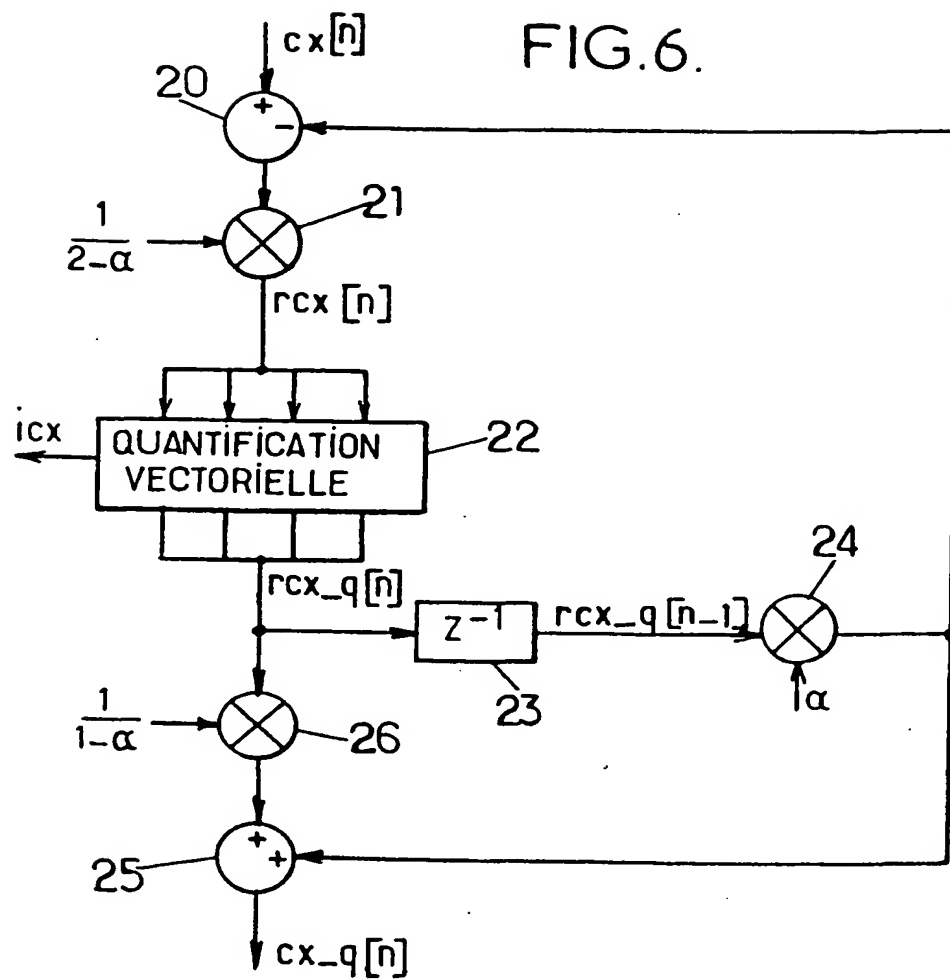


FIG. 7.

THIS PAGE BLANK (118)

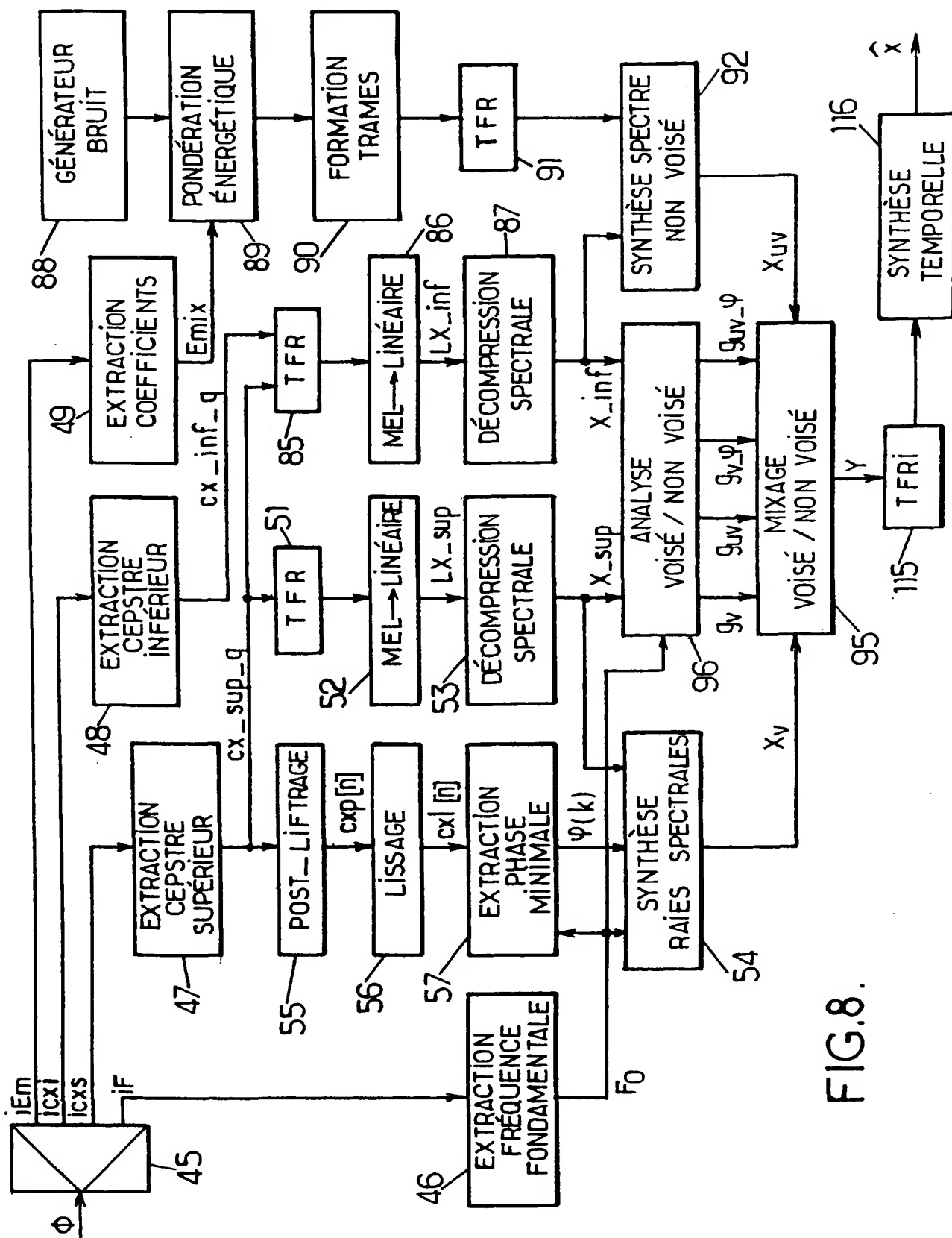
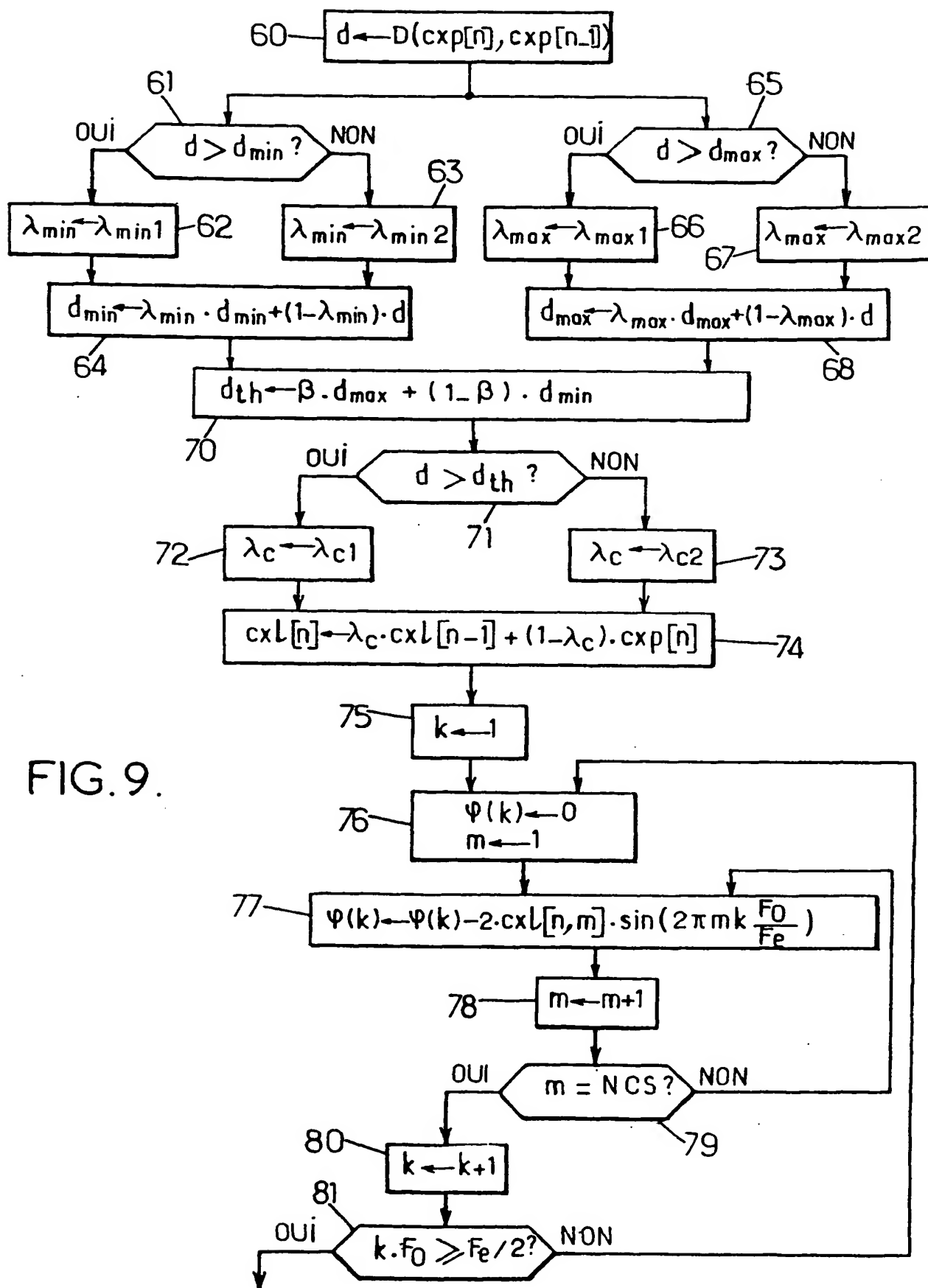


FIG. 8.

**THIS PAGE BLANK (USP:**



6/11



**THIS PAGE BLANK 0**

7/11

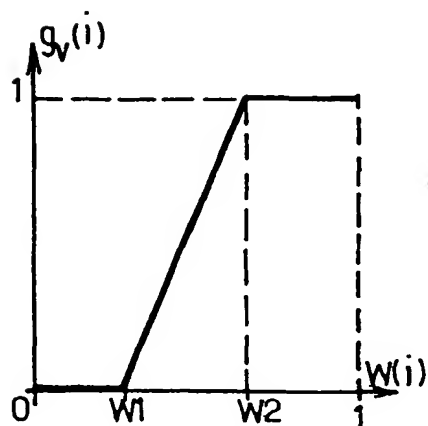
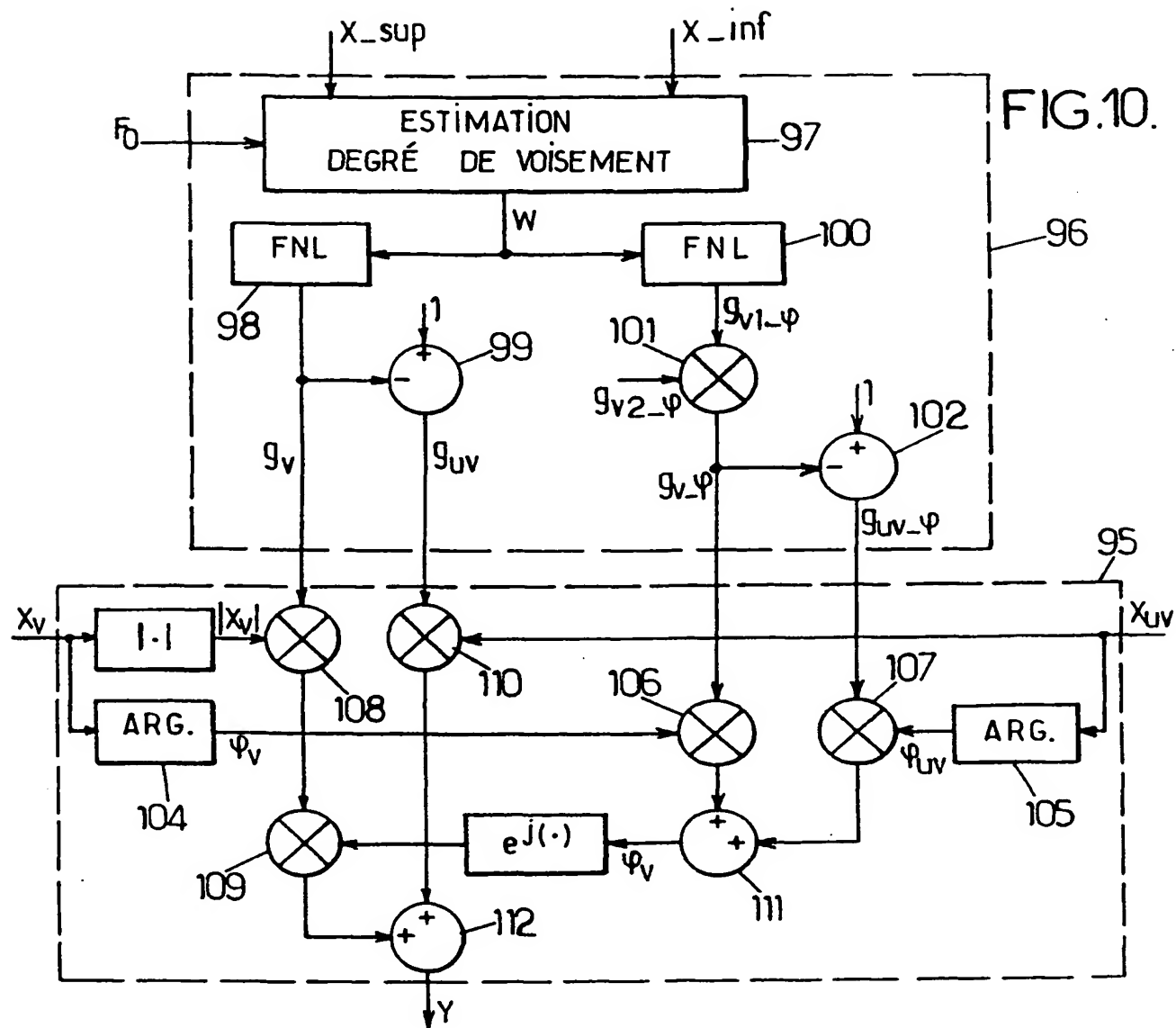


FIG.11

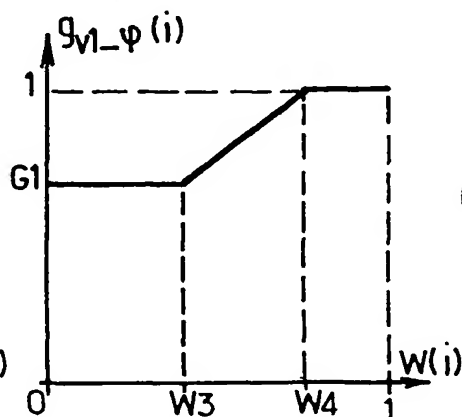


FIG.12.

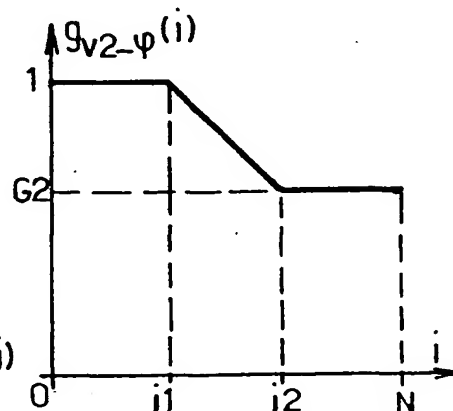


FIG.13.

THIS PAGE BLANK #15070

8/11

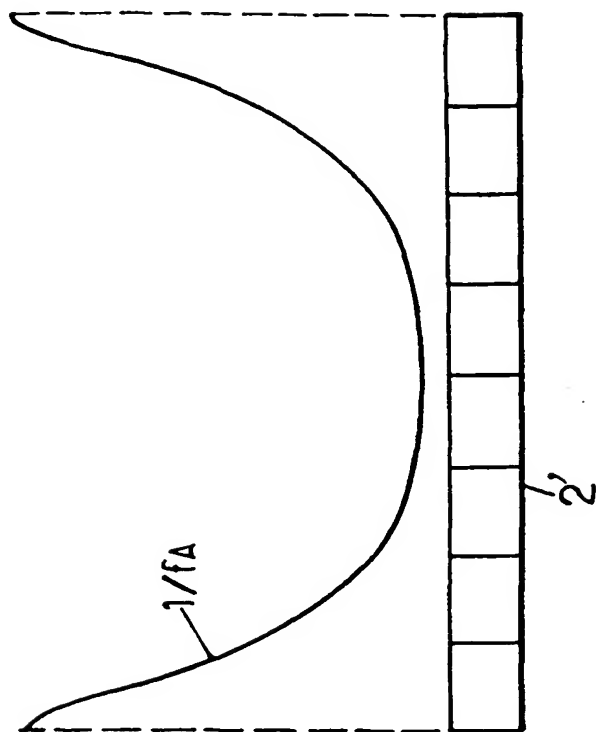


FIG. 14.

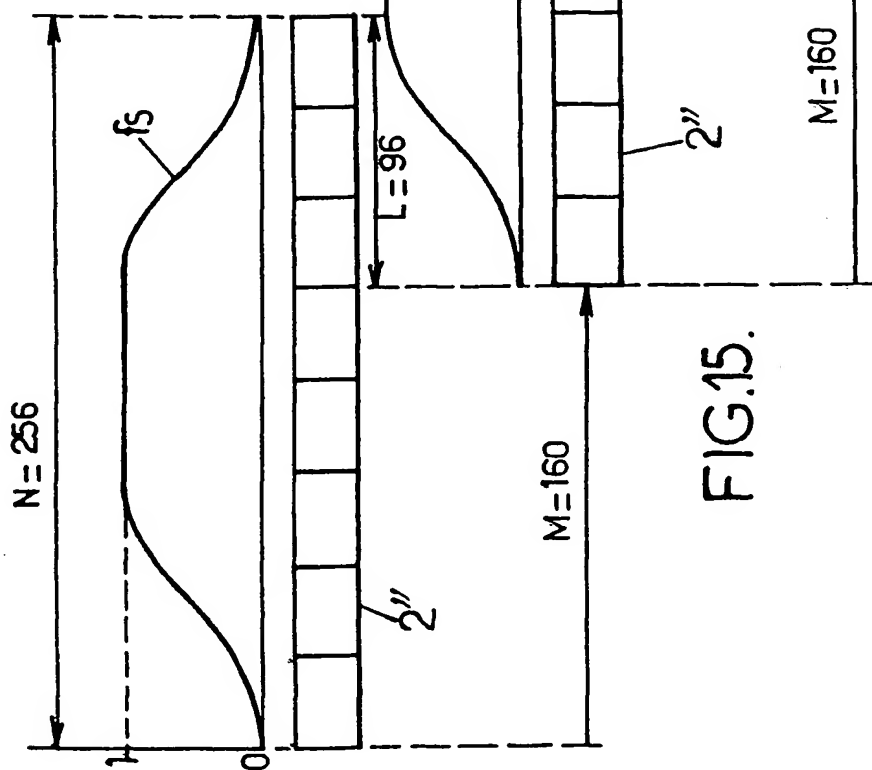


FIG. 15.

**THIS PAGE BLANK (USPTO)**

9/11

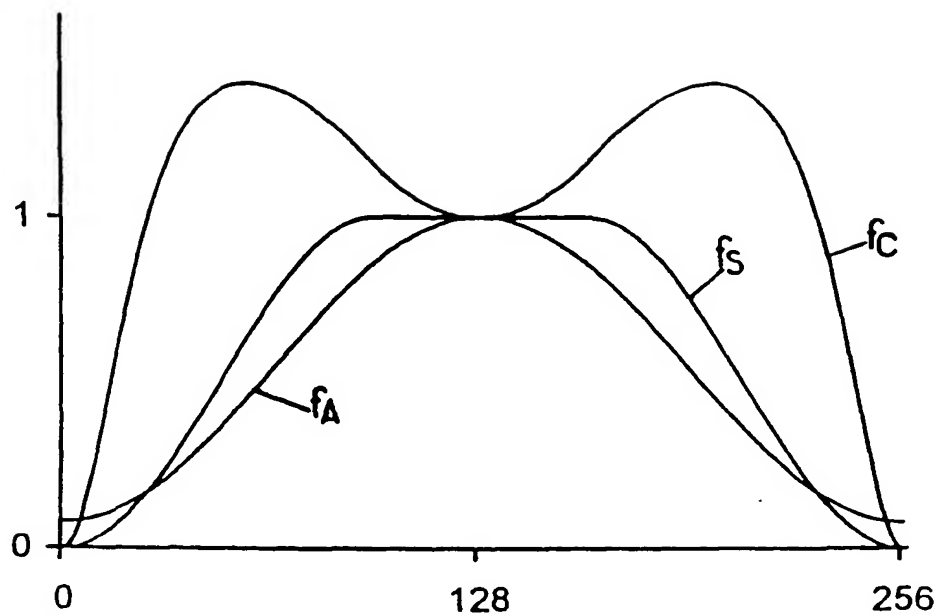


FIG.16.

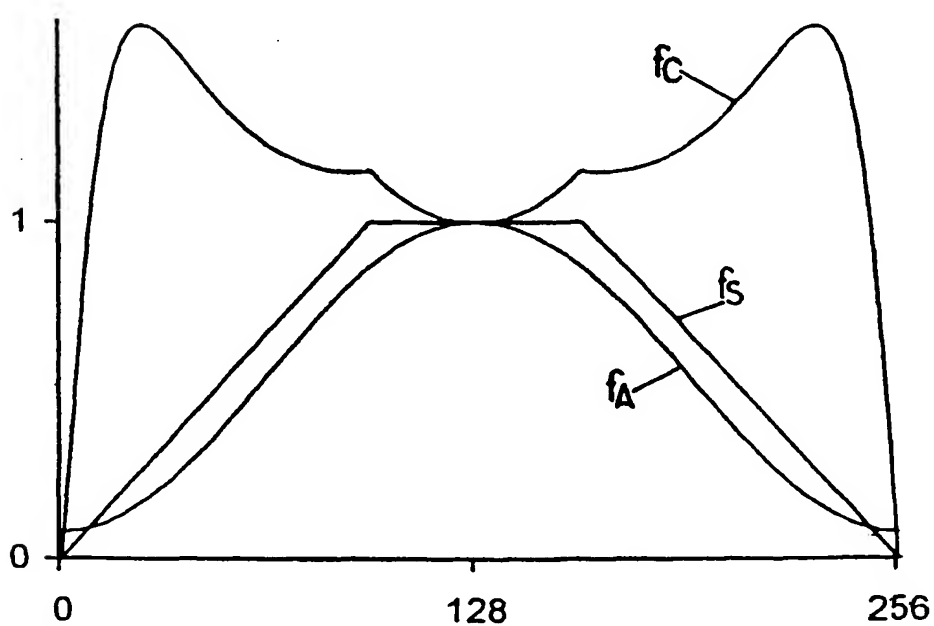


FIG.17.

THIS PAGE BLANK (USPTO)



10/11

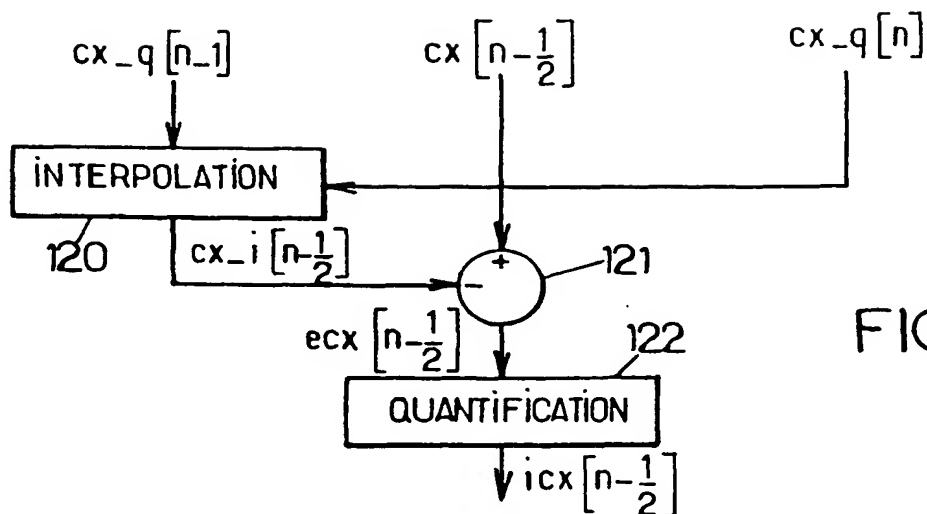


FIG.18.

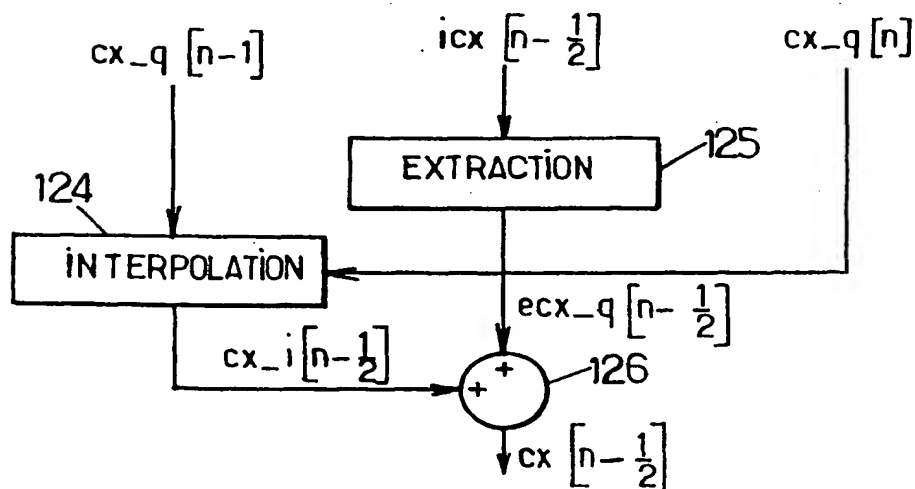


FIG.19.

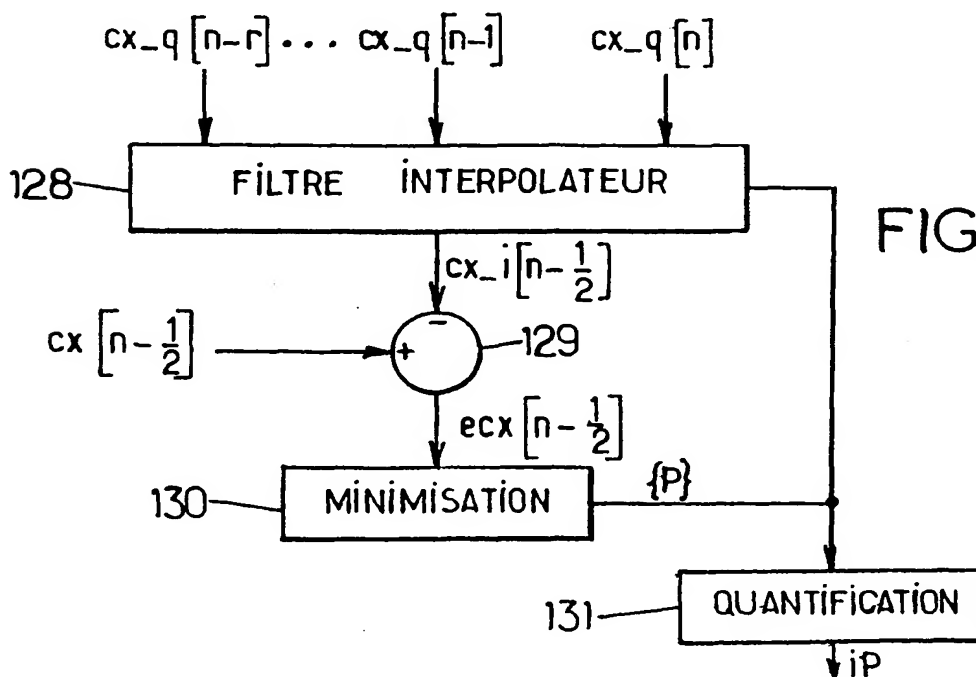


FIG.20.

THIS PAGE BLANK (USPTO)

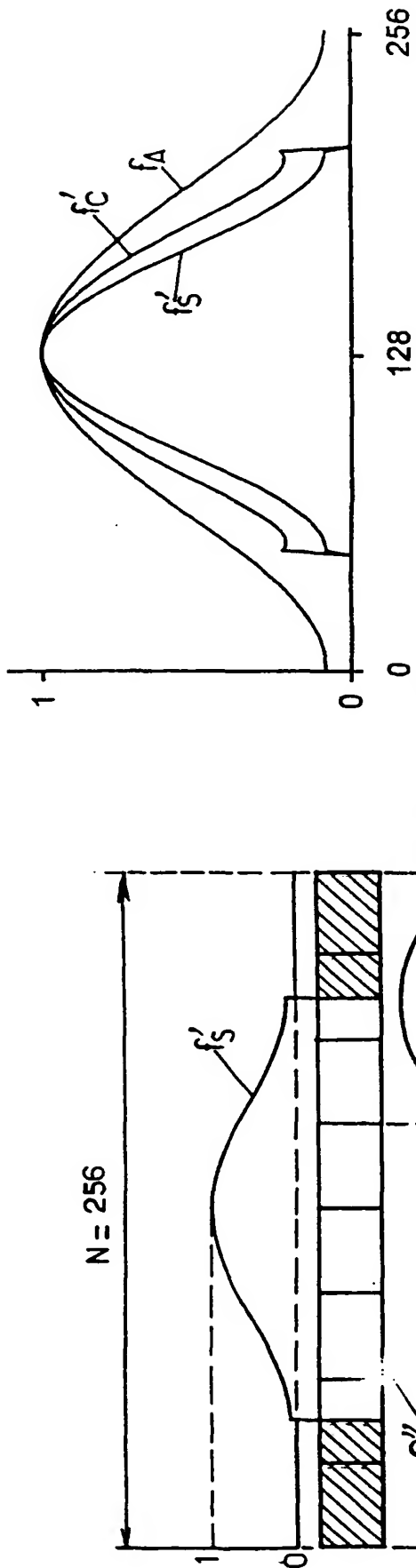


FIG. 22.

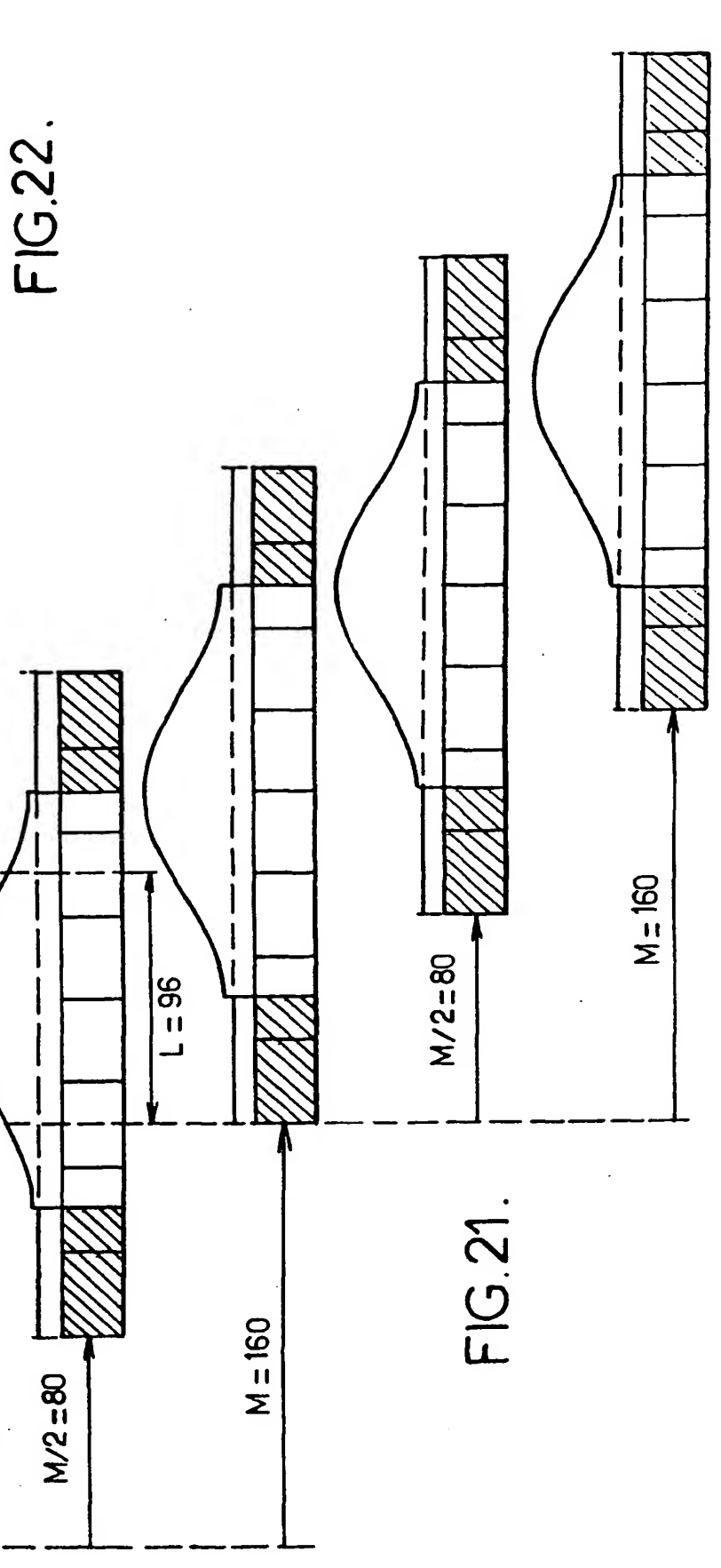


FIG. 21.

**THIS PAGE BLANK (USPTO)**